

Mind-Wandering: What Can We Learn from Eye Movements?

by

Han Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Education and Psychology)
in the University of Michigan
2020

Doctoral Committee:

Professor Kevin Miller, Chair
Professor Kai Cortina
Professor John Jonides
Professor Priti Shah

Han Zhang

hanzh@umich.edu

ORCID iD: 0000-0001-6087-0428

© Han Zhang 2020

Acknowledgments

I would like to thank my academic advisors, Kevin and Kai, for offering me the ticket to an amazing 6-year journey of learning and growth. I absolutely cannot make it without your continuous mentoring and support. Priti, I just wanted you to know that you are wonderful, both as a mentor and as a friend. You lifted me out of my darkest moments and reignited my passion for pursuing an academic career. John, I have learned so much from you in a very short period. You are the type of scientist I admire and dream to be. I sincerely appreciate you offering me a chance to work with you. I am very excited about my next chapter. Thanks to my mom and dad, for encouraging me to pursue my goals and teaching me important life lessons. Finally, I would like to thank my wife, Lizheng. You are the most important person in my life. Marrying you is the best moment I have ever had. I simply cannot ask for more. All this hard work would be meaningless without you.

Table of Contents

Acknowledgments.....	ii
List of Tables	v
List of Figures.....	vi
Abstract.....	x
Chapter 1 Introduction.....	1
1.1 What is Mind-wandering?.....	1
1.2 How to Measure Mind-wandering	5
1.3 What Happens during Mind-wandering.....	7
1.4 Using Eye-tracking to Study Mind-wandering	9
1.5 Summary and Dissertation Outline	30
Chapter 2 Mind-wandering during Visual Search.....	32
2.1 Abstract	32
2.2 Introduction	33
2.3 Methods.....	39
2.4 Results	46
2.5 Discussion	55
Chapter 3 Mind-wandering during Real-world Scene Perception.....	61
3.1 Abstract	61
3.2 Introduction	62
3.3 Method	71
3.4 Results	78
3.5 Discussion	89
Chapter 4 Mind-wandering during Reading.....	98

4.1	Abstract	98
4.2	Introduction	98
4.3	Methods	103
4.4	Results	110
4.5	Discussion	113
4.6	Pre-registered Replication	115
4.7	Additional Analysis: Mind-Wandering and Lexical Processing.....	119
4.8	General Discussion.....	120
Chapter 5	General Discussion	124
5.1	Summary of findings.....	124
5.2	Attentional decoupling during MW.	125
5.3	Task-general gaze measures of MW?	126
5.4	Concluding remarks.	144
	Bibliography	145

List of Tables

Table 1-1. Some global measures of eye movements and their definitions.....	11
Table 1-2. Global Eye Movement Features during Mind-wandering.....	16
Table 1-3. Local Eye Movement Features during Mind-wandering.....	21
Table 1-4. Pupil Size Features of Mind-wandering.....	27
Table 5-1. Global Eye Movement Measures Examined and Their Definitions.....	127
Table 5-2. Pairwise comparisons on fixation count between attentional states within each study.	129
Table 5-3. Pairwise comparisons on fixation duration between attentional states within each study.....	130
Table 5-4. Pairwise comparisons on the tau parameter between attentional states within each study.....	133
Table 5-5. Pairwise comparisons on the sum of fixation durations between attentional states within each study.	135
Table 5-6. Pairwise comparisons on the dispersion of fixations between attentional states within each study.....	136
Table 5-7. Pairwise comparisons on the percentage of area covered by fixations between attentional states within each study.....	138
Table 5-8. Pairwise comparisons on tracking loss between attentional states within each study.	140

List of Figures

Figure 1-1. Google Scholar Results for "mind-wandering" from 2005 to 2019. Source: https://csullender.com/scholar/	2
Figure 2-1. Example trials with and without the color singleton. In this figure, the target is the diamond. On 50% of the trials, one of the non-target items will appear in a different color (see right panel). The shape of the color singleton may vary from trial to trial, but it will never be the target shape. Participants' task is to indicate the line orientation inside the target item. Pilot studies suggested that the study's display setting encourages participants to make an eye movement to the target to identify the line segment. Previous studies using a similar configuration (Gaspelin et al., 2017) showed that participants were able to suppress initial eye movements to the color singleton. The two solid white circles represent the boundaries of the area of interest and were not displayed in the actual task.	42
Figure 2-2. Distractor suppression effect across all trials, as illustrated by response time (panel <i>a</i>) and initial landing position (panel <i>b</i>). Panel <i>a</i> shows that distractor-present trials were on average faster (marginally significant) compared to distractor-absent trials. Panel <i>b</i> shows that participants were less likely to have their first eye movements on the distractor than on an average non-target item in distractor-present trials. Error bars showed 95% confidence intervals.	47
Figure 2-3. Subjective reports of on-task and MW throughout the 6 experimental blocks. Error bars showed 95% confidence intervals. There were a marginally significant linear trend and a significant quadratic trend for the time course of on-task reports.	48
Figure 2-4. Response time and eye movement measures for trials reported as on-task, unintentional MW, and intentional MW. Panel <i>a</i> shows response time. Response times were overall longer during MW (intentional and unintentional combined) than during on-task episodes. Panel <i>b</i> shows the initial landing position. There was a main effect of distractor suppression (as measured by the difference between non-target and distractor for distractor-present trials), but the effect did not vary between on-task and MW trials. Panel <i>c</i> shows the time spent to locate the target. There were no significant main effects or a significant interaction. Panel <i>d</i> shows the total looking time. Participants during MW spent longer time looking at the target if a distractor was present compared to when it was absent. But this was not the case when participants were on-task. Overall, these results suggest that participants during MW could still suppress initial eye movements to the salient distractor and they seemed to be equally as fast at locating the target compared to when they were on-task. However, the presence of the distractor somehow slowed down their processing of the target, causing longer looking times and possibly slower response times.	51
Figure 2-5. Pupil size (z-scored within each participant) as a function of block number. Error bars show 95% confidence intervals.	52
Figure 2-6. The relationship between pupil size and several performance measures. Pupil sizes for each participant were divided into 10 deciles and were plotted along the X-axes. Panel <i>a</i> shows response time (z-scored) on the Y-axis. There was a U-shaped relationship for both	

distractor-present and distractor-absent trials. Panel *b*'s Y-axis shows the percentage of trials with the initial landing position on the distractor. Panel *c* shows the *z*-scores of the time needed to locate the target. Similar to response times, this measure also followed a U-shaped relationship with pupil size. Panel *d* shows the *z*-scored total looking time on the target. The measure did not follow a U-shaped relationship with pupil size. Therefore, both small and large pupil sizes were associated with reduced task performance. Moreover, the increased response time at extreme levels of pupil size cannot be attributed to increased looking time on the target. Instead, it might be related to increased time searching for the target..... 54

Figure 2-7. A boxplot showing pre-trial pupil size *z*-scores during trials classified as being on-task, unintentional MW, and intentional MW. Raw data points are plotted behind the boxplot.. 55

Figure 3-1. Example of recurrence plots. The X- and Y-axis represent the same fixation sequence, with the numbers representing the fixation index (e.g., 1st fixation, 9th fixation, etc.). Two fixations are considered recurrent (denoted as a black dot) if their physical distance is below a pre-defined threshold (usually 1 °-2 ° of visual angle corresponding to the foveal vision; Anderson et al., 2013). The red rectangle in Panel *a* illustrates that fixations 11, 19, and 20 are recurrent with fixation 1. Panel *b* illustrates determinism. As highlighted in the red rectangle, areas examined by fixations 2 and 3 are re-examined later in this trial in the same order by fixations 11 and 12. Panel *c* illustrates laminarity. As highlighted in the red rectangles, an area examined by fixation 12 is later re-examined by fixations 20 and 21, forming a vertical line. In contrast, an area initially examined by fixations 4 and 5 are later re-examined by fixation 14, forming a horizontal line. Finally, compared to panel *b*, the recurrent fixations in panel *c* are closer to the major diagonal, meaning that refixations in panel *c* occur closer together in the fixation sequence. Therefore, the center of recurrence mass (CORM) is smaller in panel *c* than in panel *b*. 65

Figure 3-2. An example trail during the study phase. A trial started with a 500-millisecond display of “Next Picture”, followed by a 1-second fixation cross. Then the picture appeared for 10 seconds, followed by a black screen for 100 milliseconds. For target trials, a thought probe would then appear to ask whether participants were mind-wandering during the picture they just saw. 75

Figure 3-3. Maps created for one of the target pictures. Panel *a* shows the original picture. Panel *b* shows the attention map generated from “on-task” fixations across participants. Panel *c* shows the attention map generated from “unintentional MW” fixations across participants. Panel *d* shows the attention map generated from “intentional MW” fixations across participants. Panel *e* shows the meaning map, following procedures in Henderson and Hayes (2017). Panel *f* shows the saliency map generated by the GBVS toolbox with default parameters..... 78

Figure 3-4. Representative scanpaths during unintentional MW and on-task episodes. Each red point represents a single fixation. Recurrent fixations are connected by a black line. As shown in the figure, the scanpath during unintentional MW (on the right) was more repetitive compared to the on-task scanpath (on the left). A high-quality version of this figure can be found at <https://osf.io/6pj9m/files/>. 81

Figure 3-5. The estimated means of eye-tracking measures during intentional MW, unintentional MW, and on-task episodes. Panels *a* to *d* show measures from Recurrence Quantification Analysis. Panels *e* and *f* show measures indicating the spatial property of fixations. Panel *g* to *i* show additional measures examined. Results are aggregated across the stimuli type. Error bars show 95% confidence intervals. The annotations indicate the significance level of the difference

between intentional/unintentional MW and the on-task episode. ** $p < .001$, * $p < .01$, * $p < .05$, † $p < .10$ 84

Figure 3-6. Correlation between meaning and saliency maps for each scene. The X-axis represents the scene number. The Y-axis represents the linear correlation between meaning and saliency maps. The horizontal dashed line represents the average value across all scenes. 85

Figure 3-7. Squared linear correlations between image properties and fixation allocation during on-task and MW episodes. Panel *a* shows relationships with meaning maps and panel *b* shows relationships with saliency maps. Error bars show 95% confidence intervals. The gray dots represent raw data points. 85

Figure 3-8. Histograms of *t*-values from paired-sample *t*-tests comparing R^2 values during MW to R^2 values calculated from subsets of on-task data with the same number of fixations. Panel *a* shows *t*-values comparing on-task and intentional MW on meaning. Panel *b* shows *t*-values comparing on-task and intentional MW on saliency. Panel *c* shows *t*-values comparing on-task and unintentional MW on meaning. Panel *d* shows *t*-values comparing on-task and unintentional MW on saliency. The vertical dashed line represents the critical *t* value in each case for declaring significance. 88

Figure 3-9. Squared semi-partial correlations of meaning and saliency maps in different attentional states. Error bars show 95% confidence intervals. The gray dots represent raw data points. 89

Figure 4-1. Mean Ratings of jokes and non-funny control sentences by Mturk workers. Error bars show the 95% Confidence Intervals. 105

Figure 4-2. (a) Recall performance and (panel *b–f*) eye movement measures, by attention (on-task, unintentional mind-wandering, intentional mind-wandering) and text type (joke, control). Error bars show the 95% confidence intervals. All measures were back-transformed to the original scale. Eye movement measures were adjusted for word length and log word frequency. 113

Figure 4-3. Fixed effects of regression analysis on the interaction between attention and word frequency. Attention (on-task, intentional mind-wandering, unintentional mind-wandering) was dummy-coded, with “on-task” as the reference level. Freq: Log10 of Word Frequency. *** $p < .001$, ** $p < .01$, * $p < .05$. Error bars show the 95% confidence intervals. 120

Figure 5-1. The number of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means. 130

Figure 5-2. The mean duration of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means. 131

Figure 5-3. The estimated Mu (μ) parameter of the distribution of fixation durations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means. 132

Figure 5-4. The estimated Sigma (σ) parameter of the distribution of fixation durations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means. 133

Figure 5-5. The estimated Tau (τ) parameter of the distribution of fixation durations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means. 134

Figure 5-6. The sum duration of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.	136
Figure 5-7. The dispersion of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.	137
Figure 5-8. The percentage of area covered by fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.	139
Figure 5-9. The number of blinks during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.	140
Figure 5-10. The percentage of tracking loss during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.	141

Abstract

Mind-wandering (MW) is ubiquitous and is associated with reduced performance across a wide range of tasks. Recent studies have shown that MW can be related to changes in gaze parameters. In this dissertation, I explored the link between eye movements and MW in three different contexts that involve complex cognitive processing: visual search, scene perception, and reading comprehension. Study 1 examined how MW affects visual search performance, particularly the ability to suppress salient but irrelevant distractors during visual search. Study 2 used a scene encoding task to study how MW affects how eye movements change over time and their relationship with scene content. Study 3 examined how MW affects readers' ability to detect semantic incongruities in the text and make necessary revisions of their understanding as they read jokes. All three studies showed that MW was associated with decreased task performance at the behavioral level (e.g., response time, recognition, and recall). Eye-tracking further showed that these behavioral costs can be traced to deficits in specific cognitive processes. The final chapter of this dissertation explored whether there are context-independent eye movement features of MW. MW manifests itself in different ways depending on task characteristics. In tasks that require extensive sampling of the stimuli (e.g., reading and scene viewing), MW was related to a global reduction in visual processing. But this was not the case for the search task, which involved speeded, simple visual processing. MW was instead related to increased looking time on the target after it was already located. MW affects the coupling between cognitive efforts and task demands, but the nature of this decoupling depends on the specific features of particular tasks.

Chapter 1 Introduction

1.1 What is Mind-wandering?

We spend a substantial amount of our waking hours wondering about things completely irrelevant to the task at hand. Bloom (1953) played audio recordings of college classes to students who had participated in them and asked them to report their thoughts at different points in the class. He found that 31% of reported thoughts during the lecture were irrelevant to the topic at hand. Klinger and Cox (1987) found that about 30% to 40% of thoughts during daily activities can be classified as mind-wandering (MW). A recent large-scale study estimated that people spent about half of their waking time mind-wandering (Killingsworth & Gilbert, 2010). The prevalence of MW is intriguing in that it highlights a distinctive aspect of human cognition that can move our focus away from the here and now. Meanwhile, the prevalence of MW is alarming because of its potential threat to task performance and even personal safety. For example, studies have shown that MW was associated with deficits in vehicle control (He, Becic, Lee, & McCarley, 2011), slower responses time to sudden events (Yanko & Spalek, 2014), and increased traffic violations and accidents (Qu et al., 2015).

MW has drawn increased attention from the scientific community in the past 15 years (as shown in Figure 1-1). A set of important discoveries has been made regarding the cognitive and neural basis of MW (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009; Kane & McVay, 2012; Schooler et al., 2011; Seli, Risko, Smilek, & Schacter, 2016), its costs and benefits (e.g., Mooneyham & Schooler, 2013; Smallwood & Andrews-Hanna, 2013; Smallwood, Fishman, &

Schooler, 2007), and the feasibility of using objective measures to detect MW (e.g., Bastian & Sackur, 2013; Bixler & D’Mello, 2016; Franklin, Smallwood, & Schooler, 2011).

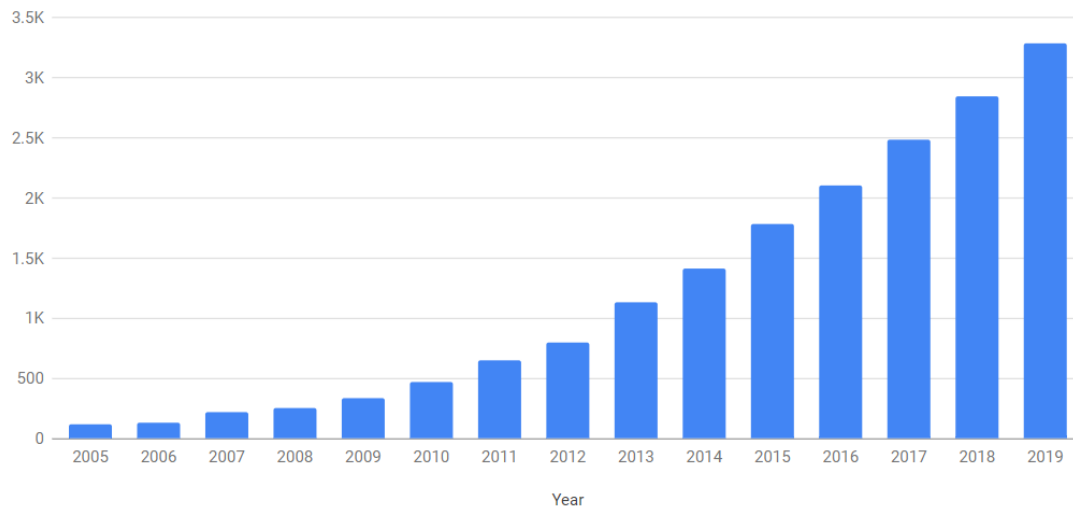


Figure 1-1. Google Scholar Results for "mind-wandering" from 2005 to 2019. Source: <https://csullender.com/scholar/>

So what is MW, exactly? The concept was proposed by Smallwood and Schooler (2006) to unify a set of somewhat related phenomena examined by earlier studies, including task-unrelated thought (Smallwood, Baracaia, Lowe, & Obonsawin, 2003), task-unrelated images and thoughts (Giambra & Grodsky, 1989), stimulus-independent thought (Antrobus, 1968), mind pops (Kvavilashvili & Mandler, 2004), zone outs (Schooler, Reichle, & Halpern, 2004), etc. Since then, there has been an ongoing debate about its definition, and there is still substantial variation in how MW is defined and measured among different research groups (Seli, Beaty, et al., 2018; Weinstein, 2018; Weinstein, De Lima, & van der Zee, 2018). The debate largely revolves around whether there is a set of defining features that distinguish MW from other types of thoughts (Christoff et al., 2018; Seli, Kane, Metzinger, et al., 2018). Some have argued that to be qualified as MW, thoughts must proceed in a relatively free, unconstrained fashion (Christoff et al., 2018). While these types of thoughts would certainly qualify as MW, the definition is not

adopted in the current dissertation for several reasons. First, this definition may be too limited in that it would exclude a lot of task-unrelated thoughts from the MW category. For example, highly constrained types of task-unrelated thoughts (e.g., deliberately planning a dinner date while sitting in a calculus class) are not considered as MW according to this definition (Christoff et al., 2018). Second, the definition (relatively free, unconstrained) is not clear enough to distinguish MW from other types of thoughts, as others have pointed out (Seli, Kane, Metzinger, et al., 2018). For example, Christoff et al. (2016) argue that MW tends to be more-deliberately constrained than dreaming but less-deliberately constrained than creative thinking and goal-directed thought. However, there are currently no valid methods to distinguish among them empirically. Third, it is unclear whether participants can report their thoughts accurately and detailed enough so that researchers can be informed about how “free and unconstrained” their thoughts are. In sum, the “free and unconstrained” thoughts might be interesting in their own right, but they can be difficult to measure and are largely orthogonal to the purpose of the current research.

In contrast with Christoff et al. (2018), Seli et al. (2018) propose a family-resemblance view of MW, in which they argue that MW consists of a family of concepts that overlap with each other but do not necessarily share a common defining feature. Instead, members of the family are organized in a graded fashion, based on how prototypical they are in the family. More prototypical members (e.g., spontaneous, task-unrelated thoughts) share more attributes with other members of the family than lower prototypical members do (e.g., meandering, unguided thoughts). This approach, according to the authors, encourages researchers to explicitly specify which member(s) in the MW family they want to investigate and explore potential differences among them. The utility of this approach can be illustrated in the recent investigation of

intentional and unintentional MW (Seli, Risko, Smilek, et al., 2016). Unintentional MW reflects a spontaneous shift from task-related to task-unrelated thoughts, despite the individual's willingness to stay on task. A perhaps less prototypical type of MW is intentional MW, a controlled and deliberate disengagement from the task. Previous research suggests that intentional and unintentional MW are dissociable. For example, increasing task difficulty reduces the rate of intentional MW but increases the rate of unintentional MW (Seli, Risko, & Smilek, 2016); task motivation and interest correlates more strongly with intentional MW than with unintentional MW (Phillips, Mills, D'Mello, & Risko, 2016; Seli, Cheyne, Xu, Purdon, & Smilek, 2015). People with ADHD, a group characterized by heightened distractibility, were reported to experience more unintentional MW but not more intentional MW (Seli, Smallwood, Cheyne, & Smilek, 2015). Trait-level intentional and unintentional MW were also found to have distinct neural correlates (Golchert et al., 2017). Therefore, while both intentional and unintentional MW belong to the MW family, they may be associated with different psychological and neurological processes.

The current dissertation adopts the family-resemblance view of MW by considering the intentionality of MW. Throughout the dissertation, being on-task is defined as “focusing on completing the task and not thinking about anything unrelated to the task”, and MW is defined as “having thoughts completely irrelevant to the current task”. Furthermore, unintentional MW is defined as “your thoughts drift away from the task despite your best intention to stay on the task” and intentional MW is defined as “you decided to think about things unrelated to the task” (Seli, Risko, & Smilek, 2016). The consistency in the definitions of intentional and unintentional MW enables me to explore potential differences between intentional and unintentional MW that are generalizable beyond a single context.

1.2 How to Measure Mind-wandering

Although MW is a common mental experience, it is a challenging topic to study scientifically. Researchers studying exogenous sources of distraction (e.g., attentional capture) can effectively manipulate attributes of exogenous stimuli and draw causal relationships between experimental manipulations and participants' responses. However, the state of MW is difficult to induce via external manipulations or infer from external behaviors. Therefore, the investigation of MW, as of now, heavily relies on participants' introspection about their internal thoughts.

One popular experience sampling method is called the "probe-caught" method (Smallwood & Schooler, 2013). Participants are intermittently interrupted and probed to report the contents of their experience during a continuous task that requires sustained attention. A short period before a positive response to the probe is considered as MW. It is worth noting that there is a large variation in the framing of thought probes, echoing the fact that different research groups conceptualized MW differently. For example, some studies simply asked participants whether or not they were mind-wandering (e.g., Franklin, Mooneyham, Baird, & Schooler, 2014). Some studies asked participants to rate their level of attention on a Likert scale (e.g., Wammes & Smilek, 2017). Participants might also be asked to differentiate between intentional and unintentional MW (e.g., Seli, Risko, Smilek, et al., 2016). Other studies asked participants to choose from several thought categories such as task-related interference, external distraction, MW, and others (e.g., Kane et al., 2017). Participants have been asked to freely describe their conscious experiences before the probe without any constraints (e.g., Stawarczyk, Cassol, & D'Argembeau, 2013). The substantial differences in probe framing pose challenges for researchers hoping to make claims of MW that are generalizable beyond a single study (Robison, Miller, & Unsworth, 2019; Weinstein, 2018; Weinstein et al., 2018).

A less popular sampling method is called the “self-caught” method (Smallwood & Schooler, 2013). In this method, participants are asked to indicate MW whenever they catch themselves doing so. A short period before the report is usually considered as being in the MW state. An advantage of the self-caught method is that there is no upper limit to the number of self-reports, thus often leading to more data and thus higher statistical power. However, the self-caught method requires participants to allocate some resources to monitor their attentional states during the ongoing task, which might be even more intrusive than thought probes. Moreover, a unique challenge for the self-caught method is to define periods during which participants are not mind-wandering, and a typical way is to use episodes without MW reports (e.g., Faber, Bixler, & D’Mello, 2018; Mills, Bixler, Wang, & D’Mello, 2015). However, it is often unclear whether participants were truly on-task during these episodes or instead failed to notice themselves mind-wandering. When probe-caught and self-caught methods are used together, self-caught reports are assumed to reflect MW with meta-awareness whereas probe-caught reports are assumed to reflect MW without meta-awareness (Reichle, Reineberg, & Schooler, 2010; Schooler et al., 2011). One study has shown that alcohol consumption increases probe-caught MW reports but lowers the incidence of self-caught reports, suggesting that self-caught reports and probe-caught reports may point to different levels or kinds of MW (Sayette, Reichle, & Schooler, 2009). However, the researcher must also consider the possibility that the increased intrusiveness when both methods are used together may severely disrupt the nature of the task.

Finally, a third way to measure MW is to use questionnaires at the end of the task (e.g., Barron, Riby, Greer, & Smallwood, 2011; Forster & Lavie, 2014; Seli, Smallwood, et al., 2015). The questionnaire method can often preserve the nature of the task and allow researchers to assess trait-level MW. However, one drawback of using questionnaires is not being able to

measure trial-level task focus, which may or may not be necessary depending on the research goal.

The subjectivity of MW reports requires the use of the triangulation strategy, whereby self-report, behavioral, and physiological measures are combined to make inferences about the underlying mental states (Smallwood & Schooler, 2015). This strategy is the underpinning of the current dissertation, in which I aim to show that subjective reports of intentional and unintentional MW are associated with both changes in behavioral measures such as response time and accuracy and changes in physiological measures such as eye movements and pupil size.

1.3 What Happens during Mind-wandering

What happens when people are mind-wandering? Smallwood and Schooler (2006; also see Schooler et al., 2011) reasoned that MW entails “attentional decoupling”, a global reduction in processing external sensory inputs, which can lead to performance deficits in tasks that require sustained external attention. Consistent with this notion, previous work has shown that MW is associated with reduced sensitivity to visual input (Barron et al., 2011; Braboszcz & Delorme, 2011; Kam et al., 2011). For example, in a Sustained Attention to Response Task (SART), Kam et al. (2011) found that MW was associated with reduced cortical processing of the target letter, as indicated by the P3 ERP component. Notably, MW was also found to be associated with reduced processing of extraneous sources of distractions. In the same experiment by Kam et al. (2011), they also found MW was associated with reduced sensory-evoked responses to a task-irrelevant shape that occasionally appeared in the task. Barron et al. (2011) found that individuals with high levels of MW in a visual oddball task showed reduced neural processing of both the rare target stimulus and the rare novel stimulus. In a sustained attention task, Esterman, Noonan, Rosenberg, & DeGutis (2013) showed reduced distractor processing during periods of low

performance (which supposedly had more MW) compared to periods of high performance. These findings seem to suggest that MW and external distractions are different mental states.

Smallwood and Schooler (2006; also see Schooler et al., 2011) proposed the attentional-resources account to explain how individuals can sustain their internal train of MW thoughts. They argue that MW “requires the coordination of information using resources under executive control” (Smallwood & Schooler, 2006, p. 549). Executive control can be described as a set of cognitive processes that regulate thoughts and behaviors for the completion of certain task goals. This theory assumes that thought processes, whether externally cued or self-generated, are served by a shared domain-general mechanism that helps to maintain the current thought processes and insulate them from any disruptions. When the mind wanders, executive control serves to maintain the continuity and integrity of internal thoughts at the cost of external processing. It is important to distinguish the attentional-resources theory from the executive-failure theory (Kane & McVay, 2012), which claims that MW is caused by failures in executive control. This account posits that sustained attention on the external task requires executive control to reduce the interference of task-unrelated events from both the external and the internal world. Evidence of the executive-failure account comes from studies looking at individual differences in MW frequency and cognitive capacities. For example, McVay and Kane (2012b) tested an individual’s working memory capacity (WMC), reading comprehension performance, and MW frequency, with each of them measured by multiple tasks. They found a positive relationship between WMC and reading performance and this relationship was partially mediated by the individual’s propensity to engage in MW. Further, Kane et al. (2007) showed that individuals with high WMC reported less MW during demanding tasks in real life, suggesting that executive control inhibits rather than supports MW. Smallwood (2013) argues that these two theories are

not mutually exclusive: the executive-failure account is about how someone gets into the state of MW, while the attentional-resources account is about what happens once someone is already in the state of MW. Specifically, the attentional-resources account is “an explanation for the continuity of self-generated thought rather than a proposal for why the episode began or is subsequently regulated” (Smallwood, 2013, p. 524).

My dissertation focuses on the nature of MW instead of individual differences in MW propensity. That is, I assume that there are common patterns across individuals when they are mind-wandering, and I seek to find these patterns. Therefore, the analyses will focus on comparing on-task vs. MW episodes across individuals. The propensity of MW for a particular individual may depend on a wide range of factors, such as their cognitive abilities and contextual factors. But these factors are not explored in the current dissertation.

1.4 Using Eye-tracking to Study Mind-wandering

The use of eye-tracking to study cognitive processes is built on the assumption that cognition actively controls eye movements (i.e., the “eye-mind” link). However, studies have also shown that our eyes may continue to move even when our attention is not fully directed towards the external world (e.g., Ehrlichman & Micic, 2012; Ehrlichman, Micic, Sousa, & Zhu, 2007; Kinsbourne, 1972; Reichle et al., 2010). Thus, the “eye-mind” link is not as tight as one might think, and MW might be one of the cases where the “eye-mind” link loosens. Studying how eye movements during these episodes differ from those during focused episodes can reveal insight into how MW affects cognitive processing. These changes in eye movements may also be useful in developing gaze-based MW detectors independent of subjective reports. Over the past few years, there is an increasing number of studies investigating the link between eye movement

and MW. But before diving into these studies, it might be worthwhile to introduce several concepts that will be helpful when interpreting the results.

1.4.1 A brief introduction to eye-tracking.

Our eyes are an important tool for processing the external world. However, our visual system is rather limited in that the clearest vision is restricted to the fovea, a tiny pit on our retina. As a result, we must move our eyes to different locations to facilitate visual processing in complex tasks such as reading and scene perception. The physical trajectories traversed by eye movements can be highly informative of the viewer's moment-to-moment cognitive activities (Henderson, 2003; Land & Hayhoe, 2001; Rayner, 1998). Eye movements primarily consist of *fixations* and *saccades*. During a fixation, our eyes remain relatively still at one location. The fixated location receives the highest visual acuity, allowing visual input to occur. The duration of a fixation is a rough measure of the time needed to sufficiently process information at one location. For example, during reading comprehension, low-frequency words usually receive longer looking times than high-frequency words (e.g., Kliegl, Grabner, Rolfs, & Engbert, 2004). Saccades are rapid movements of the eyes that travel from one place to another, which usually marks the transition between two consecutive fixations. Compared to fixations, the duration of a saccade is much shorter (e.g., 225 ms vs. 30 ms in reading; Rayner, 1998) and visual input is largely suppressed during a saccade. An important property of a saccade is its direction, which can indicate the location from which the viewer wishes to seek information. For example, about 10% - 15% of saccades during English reading go back to previously inspected words (Rayner, 1998). These so-called "regressions" indicate the need for re-processing previous information and play important roles in reading comprehension (Schotter, Tran, & Rayner, 2014).

Faber et al. (2018) as well as others (e.g, Bixler & D’Mello, 2016; Hutt et al., 2017; Mills et al., 2015) broadly categorize eye movement measures into two categories: global measures and local measures. Global measures are those calculated without resorting to the specific content of the stimuli, such as the total number of fixations/saccades, the average duration of fixations/saccades, the length of saccades, the angle of saccades, etc. Some global measures and their definitions are listed in Table 1-1. Local measures are just the opposite in that they are related to the stimuli content. For example, local measures may include the amount of time to locate a specific target, the total looking time in a particular region, the number of saccades going back to previously inspected words, etc. Faber et al. (2018) argue that global measures, compared to local measures, are more robust against missing, poor, or invalid gaze data, which are all likely scenarios outside of laboratory settings. Therefore, gaze-based detectors based on global measures may be a more ecologically valid option that can be applied to different tasks and tracking scenarios. However, being task-independent implies that global measures often provide limited information about specific cognitive processes in a specific task scenario. Thus, if the researcher’s goal is to understand how MW might disrupt specific cognitive processes, a detailed analysis of local features is often necessary. Of course, global and local measures are sometimes correlated (e.g., the total number of fixations vs. the number of fixations in a particular region). These two set of measures are often complementary to each other and tell a coherent story. To fully understand the link between MW and eye movements, both types of measures should not be ignored.

Table 1-1. Some global measures of eye movements and their definitions.

Global Measures	Definitions
-----------------	-------------

Fixation Count	The total number of fixations.
Fixation Duration Mean	The average duration of fixations.
Fixation Duration Sum	The duration sum of fixations.
Fixation Dispersion	The root mean square of the Euclidean distance from each fixation to the average position of all fixations. It is reported on a 0-1 scale by normalizing on the maximum dispersion possible, with higher values indicating greater dispersion.
Area Covered	The total area covered by fixations (a circular region with a radius of 60 pixels) in proportion to the screen size.
Blink Count	The number of blinks.
Tracking Loss	The percentage of gaze signal not recorded by the eye-tracker.

Because eye movements naturally unfold over time, the sequential aspect of eye movements (i.e., scanpaths) can also be informative of the viewer’s cognitive processing. For example, early work by Yarbus (1967) showed that eye movements on the same painting differed depending on task instructions. Noton and Stark (1971a, 1971b) further proposed that the sequential information of eye movements is an external reflection of the viewer’s internal cognitive representation. There is an entire field devoted to extracting meaningful information from highly complex scanpath patterns (for a review, see Anderson, Anderson, Kingstone, & Bischof, 2015). A growing body of literature has employed scanpath analysis to study online cognitive processing during scene perception (Foulsham & Underwood, 2008), visual search (Dewhurst et al., 2018), reading (von der Malsburg & Vasishth, 2013), social communication (Richardson & Dale, 2005), domain expertise (McIntyre & Foulsham, 2018), etc.

Besides eye movements, miscellaneous measures captured by eye-trackers, such as pupil size and eye blinking, can also indicate the viewer's cognitive processing. Recent studies have linked these measures to MW (e.g., Konishi, Brown, Battaglini, & Smallwood, 2017; Smilek, Carriere, & Cheyne, 2010; Unsworth & Robison, 2016). Their relationships will be discussed in more detail in section 1.4.4.

Eye movements are highly complex, and this complexity creates challenges for researchers who wish to extract useful information from them. There are two general ways to deal with this complexity, as evident in the previous literature on eye movements during MW. The first is a hypothesis-driven approach. Studies using this approach would usually specify a limited set of measures that are most relevant to their hypotheses and see if the results confirm or reject their hypotheses (e.g., Foulsham, Farley, & Kingstone, 2013; Reichle et al., 2010; Smilek et al., 2010; Steindorf & Rummel, 2019; Unsworth & Robison, 2016). The hypothesis-driven approach cares about how the results help to advance our understanding of the underlying mechanisms of MW. But sometimes it is difficult to justify why certain measures are chosen but not others, and there is a danger of “cherry-picking” measures that are significant. A strong theory or a pre-registered plan can help to alleviate these problems. The second is a data-driven approach. These studies typically seek to categorize MW and on-task self-reports based on gaze data (e.g., Bixler & D’Mello, 2016; Faber et al., 2018; Hutt et al., 2017; Hutt, Mills, White, Donnelly, & D’Mello, 2016; Mills et al., 2015). Studies using this approach would typically compute a large number of measures, of which many can be deemed as “unconventional” (e.g., the minimum, max, and range of fixation durations, besides the more “conventional” mean value). Then, they would try many combinations between different measures, different classifiers, and different parameter settings to find the situation in which the classifier performs

the best. Because this approach is essentially blind to theory, it is often not clear how certain results can advance our understanding of the mechanisms of MW (e.g., MW had a smaller saccade angle maximum value compared to being on-task; Hutt et al., 2016). Moreover, it is often unclear whether the same findings would hold for unseen data.

In the next three sections, I summarize and analyze previous findings on eye movement correlates of MW. Section 1.4.2 covers global measures of MW; section 1.4.3 covers local measures of MW; section 1.4.4 covers pupil size during MW.

1.4.2 Global eye movement features of MW.

Table 1-2 summarizes the results of previous studies using global measures. Because the test between MW and on-task episodes involves within-subject comparisons, Cohen's d_z was used to indicate the strength of the difference. A positive value indicates that the measure was greater during MW than during on-task episodes; Values were obtained from the original papers when possible. There were multiple cases where d_z was not provided in the paper: in some cases, a paired-sample t -test was conducted but the effect size was not reported; in other cases, the design was more complex than a paired-sample t -test and other test statistics were reported. In this case, translating the effect size to d_z is desirable because it allows for a straightforward interpretation across different designs. Therefore, in case d_z was not reported, it was either estimated from test statistics such as t and F values or the means and the standard deviations provided in the paper. Also, note that this table does not list all global measures examined in the previous literature. A measure was listed here either because (1) its Cohen's d_z absolute value was greater than .20 or (2) more than one study examined this measure. These criteria excluded many measures that were examined in a single study or had low effect sizes.

Several generalizations can be made from these results. First, MW is generally associated with fewer fixations and fewer saccades. Therefore, there seems to be less visual processing in general during MW compared to being on-task, a pattern consistent with the idea of attentional decoupling. Second, the duration of fixations seems to be longer during MW compared to being on-task. As said previously, the duration of a fixation roughly indicates the amount of time needed to sufficiently process information in a certain region. The increased fixation may suggest that visual processing is less efficient at fixated locations during MW. Third, fixations seem to be located further away from each other during MW, as indicated by increased fixation dispersion. Relatedly, saccades seem to take longer duration MW, which can be caused by increased travel distance.

The rate of blinking is associated with time-on-task and fatigue (Stern, Boyer, & Schroeder, 1994a, 1994b; Maffei & Angrilli, 2018). Smilek et al. (2010) showed that MW was also associated with increased blinks in a reading task. The authors suggest that eye blinks are a physical embodiment of the attentional decoupling process such that the closing of the eyelids helps to insulate internal thoughts from the visual input. However, subsequent studies obtained mixed results. Even for those who did obtain significant results, the effect size seems much smaller compared to the original study (e.g., Krasich et al., 2018, exp. 1). These discrepancies may partly arise from how blinks are detected. Smilek et al. (2010) defined blinks as any periods with more than 2 ms of tracking loss. However, there are many reasons besides blinks that can cause tracking loss, such as head movement, reduction in pupil size, microsleeps, looking outside of the screen, etc., all of which may tend to occur during MW. Using blink-detection algorithms that take into account the rapid changes in pupil size before and after blinks (e.g., Hershman,

Henik, & Cohen, 2018), or using event-related potentials to record and detect blink artifacts may help to tease apart these confounds.

Table 1-2. Global Eye Movement Features during Mind-wandering.

Measure	Cohen's d_z (MW – On-task)	Task	Method	N	N of probes	Paper
Blink Count	(+)	Reading	Probe	22	10	(Uzzaman & Joordens, 2011)
	1.23 ^a	Reading	Probe	12	20	(Smilek et al., 2010)
	0	Reading	Self	178		(Faber et al., 2018)
	0.06 ^b	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
	0.13 ^b	Scene Viewing	Probe	51	8	(Krasich et al., 2018, Exp. 1)
	0.08 ^b	Scene Viewing	Probe	41	3	(Krasich et al., 2018, Exp. 2)
Blink Duration Mean	0.54 ^b	Breath Count	Self	2		(Grandchamp, Braboszcz, & Delorme, 2014)
Blink Rate	0.39 ^b	Breath Count	Self	2		(Grandchamp et al., 2014)
Fixation Count	0.14 ^b	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
	0.35 ^a	Reading	Probe	26	9	(Foulsham et al., 2013)

	-0.41	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
	-0.43 ^b	Scene Viewing	Probe	51	8	(Krasich et al., 2018, Exp. 1)
	-0.51 ^b	Scene Viewing	Probe	41	3	(Krasich et al., 2018, Exp. 2)
	-0.033	Reading	Self	178		(Faber et al., 2018)
Fixation Dispersion Mean	0.13 ^b	Scene Viewing	Probe	51	8	(Krasich et al., 2018, Exp. 1)
	0.23	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
	0.53 ^b	Scene Viewing	Probe	41	3	(Krasich et al., 2018, Exp. 2)
	0.03	Reading	Self	178		(Faber et al., 2018)
Fixation Duration Kurtosis	-0.58 ^b	Reading	Self	30		(Bixler & D'Mello, 2015)
	0.11 ^b	Scene Viewing	Probe	51	8	(Krasich et al., 2018, Exp. 1)
Fixation Duration Mean	0.35 ^b	Scene Viewing	Probe	12	20	(Krasich et al., 2018, Exp. 2)
	0.33 ^a	Reading	Probe	41	3	(Smilek et al., 2010)

	0.41 ^a	Reading	Probe	26	9	(Foulsham et al., 2013)
Fixation Duration Median	0.223	Reading	Self	178		(Faber et al., 2018)
Fixation Duration Skew	-0.27	Virtual Tutor	Probe	105	10	(Hutt et al., 2016)
	-0.53 ^b	Reading	Self	30		(Bixler & D'Mello, 2015)
Horizontal Gaze Position SD	-0.97 ^a	Simulated Driving	Self	18		(He et al., 2011)
Horizontal Saccade %	-0.312	Reading	Self	178		(Faber et al., 2018)
Saccade Angle Max	-0.24	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
Saccade Angle Max (Absolute)	-0.186	Reading	Self	178		(Faber et al., 2018)
Saccade Angle Max (Relative)	-0.133	Reading	Self	178		(Faber et al., 2018)
Saccade Angle Range	-0.26	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
Saccade Angle Range (Relative)	-0.204	Reading	Self	178		(Faber et al., 2018)
Saccade Count	-0.218	Reading	Self	178		(Faber et al., 2018)

	-0.31	Film Comprehension	Self	60		(Mills et al., 2015)
	-0.4	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
Saccade Duration Max	0.109	Reading	Self	178		(Faber et al., 2018)
	0.23	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
Saccade Duration Mean	0.24	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
	0.32	Virtual Tutor	Probe	105	10	(Hutt et al., 2016)
Saccade Duration Median	0.22	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
Saccade Velocity Mean	-0.32	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
Saccade Velocity Median	-0.33	Virtual Tutor	Probe	135	12	(Hutt et al., 2017)
Smooth Pursuit Duration Range	0.3	Film Comprehension	Self	60		(Mills et al., 2015)
Smooth Pursuit Velocity SD	-0.28	Film Comprehension	Self	60		(Mills et al., 2015)

Notes.

1. This table does not list all global measures examined in the previous literature. A measure was listed here either because (1) its Cohen's d_z absolute value was greater than .20 or (2) more than one study examined this measure. These criteria excluded many measures that were examined in a single study or had low effect sizes.
2. Cohen's d_z : Positive values indicate MW > On-task; Negative values indicate MW < On-task. Values were obtained from the original papers when possible. (+) and (-) only indicates the direction of the difference. Superscript ^a indicates that the value was calculated from test statistics (e.g., t and F); superscript ^b indicates that the value was calculated from means and standard deviations provided in the paper (assuming within-subject correlation is .5).

-
3. Method: “Probe” indicates studies that used the probe-caught method; “Self” indicates studies that used the self-caught method.

1.4.3 Local eye movement features of MW.

Table 1-3 summarizes local eye movement measures of MW. As evident in the table, most of these measures were tested in the context of reading comprehension. The stimuli in reading tasks (i.e., words) often have clear boundaries, thus allowing researchers to easily define local measures based on individual words. Some commonly used local measures include first-fixation duration (the duration of the first fixation on a word during first-pass reading), gaze duration (the duration of all fixations on a word during first-pass reading), single fixation duration (the looking time on a word with only one fixation), and total looking time (the duration of all fixations on a word). Other measures may include inter-word regressions and word skipping. A review of eye movement measures during reading is beyond the focus of this chapter and can be found elsewhere (for more information, see Kliegl et al., 2004; Rayner, 1998; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

From Table 1-3, we can see that the duration measures (first-fixation duration, gaze duration, total looking time, etc.) tend to be longer during MW. This pattern echoes the global pattern that fixations tend to last longer during MW (see Table 1-2) and may suggest that readers need a longer time to sufficiently process words that are fixated during MW. Interestingly, readers also tend to fixate less on words but more on off-text regions during MW. This finding suggests that task-relevant processing is reduced during MW.

One of the classical effects in the reading literature is the word frequency effect – the fact that low-frequency words tend to receive longer looking time compared to high-frequency words. This effect suggests that top-down cognition actively controls eye movements during reading (Pollatsek, Reichle, & Rayner, 2006). Reichle et al. (2010) showed that the word frequency effect was absent during MW. Subsequent studies obtained somewhat similar results (Foulsham et al., 2013; Steindorf & Rummel, 2019). These results are globally consistent with the attentional decoupling account by showing a decoupling between cognition and eye movements during MW. However, it should be noted that there are inconsistencies in the measures that these studies reported. For example, while Foulsham et al. (2013) found that the word frequency effect on gaze duration was significantly reduced during MW, Steindorf & Rummel (2019) did not find the same effect with the same measure but instead found a similar pattern for fixation count. Also note that a similar pattern was not found for word length (Frank et al., 2015), a variable highly correlated with word frequency.

Table 1-3. Local Eye Movement Features during Mind-wandering.

Measure	Cohen's <i>d</i> (MW - On-task)	Task	Method	N	N of probes	Paper
	(+)	Reading	Probe	4	152	(Reichle et al., 2010)
First-fixation Duration Mean	$\geq 1.175^a$	Reading	Self	4		(Reichle et al., 2010)
	0.23 ^b	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
First-pass Duration Mean	0.21	Reading	Probe	178	28	(Bixler & D'Mello, 2016)

First-pass Duration SD	0.2	Reading	Probe	178	28	(Bixler & D'Mello, 2016)
	-3.44 ^a	Reading	Self	4		(Reichle et al., 2010)
First-pass Fixation Count	-0.1 ^a	Reading	Probe	29	20	(Frank, Nara, Zavagnin, Touron, & Kane, 2015)
	(-)	Reading	Probe	4	152	(Reichle et al., 2010)
	0.21	Reading	Probe	178	28	(Bixler & D'Mello, 2016)
	2.17 ^a	Reading	Probe	4	152	(Reichle et al., 2010)
Gaze Duration Mean	$\geq 1.175^a$	Reading	Self	4		(Reichle et al., 2010)
	0.02 ^a	Reading	Probe	29	20	(Frank et al., 2015)
Gaze Duration SD	0.24	Reading	Probe	178	28	(Bixler & D'Mello, 2016)
Interword Regression %	-0.22 ^a	Reading	Probe	26	9	(Foulsham et al., 2013)
	(-)	Reading	Probe	4	152	(Reichle et al., 2010)
Interword Regression Count	-4.66 ^a	Reading	Self	4		(Reichle et al., 2010)

	0.4 ^a	Reading	Probe	29	20	(Frank et al., 2015)
	> -.004 ^b	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
Line Cross Saccades %	0.47 ^b	Reading	Self	30		Bixler & D'Mello (2015)
Number of Saccades away from Saliency AOI	-0.39	Film Comprehension	Self	60		(Mills et al., 2015)
Number of Saccades nearly onto Saliency AOI	-0.35	Film Comprehension	Self	60		(Mills et al., 2015)
	2.36 ^a	Reading	Self	4		(Reichle et al., 2010)
Off-text Fixation Count	0.29 ^a	Reading	Probe	29	20	(Frank et al., 2015)
	1.86 ^a	Reading	Probe	4	152	(Reichle et al., 2010)
Percent of Gaze Dwell Time on Side Mirrors	-0.75 ^a	Simulated Driving	Self	18		(He et al., 2011)
	0.39 ^a	Reading	Probe	26	9	(Foulsham et al., 2013)
Reading Time	0.12 ^b	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
Reading Time Ratio	0.22	Reading	Probe	178	28	(Bixler & D'Mello, 2016)

Regression Fixations %	0.54 ^b	Reading	Self	30		(Bixler & D'Mello, 2015)
Single Duration Mean	0.2	Reading	Probe	178	28	(Bixler & D'Mello, 2016)
Single Duration SD	0.22	Reading	Probe	178	28	(Bixler & D'Mello, 2016)
Smooth Pursuit with Balloon AOI (frames)	-0.37	Film Comprehension	Self	60		(Mills et al., 2015)
Smooth Pursuit within 2° Saliency AOI (frames)	-3.8	Film Comprehension	Self	60		(Mills et al., 2015)
Time Off Text	0.34 ^a	Reading	Probe	29	20	(Frank et al., 2015)
Total Looking Time Mean	2.06 ^a	Reading	Probe	4	152	(Reichle et al., 2010)
	$\geq 1.175^a$	Reading	Self	4		(Reichle et al., 2010)
Total Run Count	-0.47 ^a	Reading	Probe	22	10	(Uzzaman & Joordens, 2011)
Within-word Regression Count	-0.56 ^a	Reading	Probe	22	10	(Uzzaman & Joordens, 2011)
	-0.02 ^b	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
Word Frequency Effect on Regressions into Target Word	-0.14 ^a	Reading	Probe	107	10	(Steindorf & Rummel, 2019)

Word Frequency Effect on Fixation Count	-0.25 ^a	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
Word Frequency Effect on Gaze Duration	-0.28 ^a	Reading	Probe	26	9	(Foulsham et al., 2013)
	-0.09 ^a	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
Word Frequency Effect on Total Looking Time	-0.15 ^a	Reading	Probe	107	10	(Steindorf & Rummel, 2019)
Word Length Effect on Gaze Duration	-0.034 ^a	Reading	Probe	29	20	(Frank et al., 2015)
Words Fixated	(-)	Reading	Probe	4	152	(Reichle et al., 2010)
	-2.7 ^a	Reading	Self	4		(Reichle et al., 2010)
	-0.84 ^a	Reading	Probe	12	20	(Smilek et al., 2010)
	-0.22 ^a	Reading	Probe	29	20	(Frank et al., 2015)
Words Skipped	0.2	Reading	Probe	178	28	(Bixler & D'Mello, 2016)

Notes.

1. This table does not list all local measures examined in the previous literature. A measure was listed here either because (1) its Cohen's d_z absolute value was greater than .20 or (2) more than one study examined this measure. These criteria excluded many measures that were examined in a single study or had low effect sizes.
2. Cohen's d_z : Positive values indicate MW > On-task; Negative values indicate MW < On-task. Values were obtained from the original papers when possible. (+) and (-) only indicates the direction of the difference. Superscript ^a indicates that the value was calculated from test statistics (e.g., t and F); superscript ^b indicates that the value was calculated from means and standard deviations provided in the paper (assuming within-subject correlation is .5).
3. Method: "Probe" indicates studies that used the probe-caught method; "Self" indicates studies that used the self-caught method.

1.4.4 Pupil size during MW.

Pupil size was used in early studies as an indicator of cognitive load (e.g., Kahneman & Beatty, 1966). Interests in pupil size recently re-emerged, thanks to the *adaptive-gain theory* (Aston-Jones & Cohen, 2005) that links pupil size to activities in the locus coeruleus (LC). The LC is responsible for the synthesis of norepinephrine (NE) and has widespread projections throughout the neocortex. The LC-NE system modulates arousal, attention, and responses to stress; deficits in the LC-NE system has been linked to anxiety, depression, and attention deficits hyperactivity disorder (Benarroch, 2009). According to Aston-Jones and Cohen (2005), there is an inverted-U relationship between tonic LC activity and performance on tasks that require focused attention, resembling the Yerkes-Dodson Law (Yerkes & Dodson, 1908). Specifically, the optimal level of performance is achieved with an intermediate level of tonic LC activity, during which individuals maintain focus on the current task (exploitation mode); too high levels of tonic LC activity leads to a distractible attentional state (exploration mode), whereas too low levels of tonic LC activity leads to sleep and drowsiness. Importantly, there is a correlation between tonic pupil diameter and LC activity that exists in both monkeys (Rajkowski, Kubiak, & Aston-Jones, 1993) and humans (Gilzenrat, Cohen, Rajkowski, & Aston-Jones, 2003). Therefore, tonic pupil diameter has been thought to indirectly reflect activities of the LC-NE system and thus be associated with task performance and attentional state.

Consistent with the adaptive-gain theory, previous studies show that performance is lower when the pupil size is very small or very large (Gilzenrat, Nieuwenhuis, & Cohen, 2010; Kristjansson, Stern, Brown, & Rohrbaugh, 2009; Murphy, Robertson, Balsters, & O'connell, 2011; van den Brink, Murphy, & Nieuwenhuis, 2016). Relatedly, MW was found to be associated with both smaller and larger pupil sizes compared to on-task episodes, and sometimes

MW was not associated with changes in pupil size at all (see Table 1-4). These findings are broadly consistent with the notion that MW is a heterogeneous term and may occur under both high and low arousal states. While there are inconsistencies regarding the tonic pupil size, MW was consistently shown to be associated with reduced task-evoked pupillary responses (phasic activity). This is consistent with the adaptive-gain theory because both overly low and overly high arousal states reduce the phasic LC activation.

Table 1-4. Pupil Size Features of Mind-wandering.

Measure	Cohen's d_z (MW - On-task)	Task	Method	N	N of probes	Paper
	0.54 ^a	SART	Probe	27	50	(Jubera-García, Gevers, & Van Opstal, 2019, Exp. 1)
	0.08 ^a	SART	Probe	33	50	(Jubera-García et al., 2019, Exp. 2)
	-2.19 ^a	Stop-signal	Probe	20	40	(Mittner et al., 2014)
Pretrial Pupil Size Mean	-0.32 ^a	Psychomotor Vigilance Task	Probe	37	30	(Unsworth & Robison, 2016)
	-0.3	Psychomotor Vigilance Task	Probe	152	15	(Unsworth & Robison, 2018, Exp. 1)
	-0.32	Stroop	Probe	157	12	(Unsworth & Robison, 2018, Exp. 1)
	-0.09	Psychomotor Vigilance Task (with an internal secondary task)	Probe	28	30	(Unsworth & Robison, 2018, Exp. 2)

	-0.48	Psychomotor Vigilance Task (Varied Pace)	Probe	35	30	(Unsworth & Robison, 2018, Exp. 3)
	-0.01	Psychomotor Vigilance Task (Fixed 5s Pace)	Probe	36	30	(Unsworth & Robison, 2018, Exp. 3)
	-0.52	Psychomotor Vigilance Task (Fixed 2s Pace)	Probe	39	23	(Unsworth & Robison, 2018, Exp. 4)
	0.13	Psychomotor Vigilance Task (Fixed 8s Pace)	Probe	42	23	(Unsworth & Robison, 2018, Exp. 4)
	0	Whole-report Task	Probe	36	8	(Robison & Unsworth, 2019, Exp. 2)
	-1.15 ^b	Breath Count	Self	2		(Grandchamp et al., 2014)
	0.73 ^a	Self-paced Reading	Probe	13	26	(Franklin, 2013)
Pupil Size Mean	-1.26 ^a	0-back and 1-back	Probe	36	19	(Konishi et al., 2017)
	0.1 ^a	Vigilance Task with Distraction Words	Probe	42	28	(Pelagatti, Binda, & Vannucci, 2018)
	0.24	Reading	Probe	178	28	(Bixler & D'Mello, 2016)
Pupil Size SD	-0.03	Reading	Self	178		(Faber et al., 2018)
	-0.1 ^b	Breath Count	Self	2		(Grandchamp et al., 2014)

Pupil Size Skew	-0.222	Reading	Self	178		(Faber et al., 2018)
Pupil Size Change Rate	1.00 ^b	Breath Count	Self	2		(Grandchamp et al., 2014)
Task Evoked Pupil Response (Peak - Baseline)	-0.62 ^a	SART	Probe	27	50	(Jubera-García et al., 2019, Exp. 1)
	-1.59 ^a	SART	Probe	33	50	(Jubera-García et al., 2019, Exp. 2)
	-0.44 ^a	Psychomotor Vigilance Task	Probe	37	30	(Unsworth & Robison, 2016)
	-0.28	Psychomotor Vigilance Task	Probe	152	15	(Unsworth & Robison, 2018, Exp. 1)
	-0.33	Stroop	Probe	157	12	(Unsworth & Robison, 2018, Exp. 1)
	-0.59	Psychomotor Vigilance Task	Probe	28	30	(Unsworth & Robison, 2018, Exp. 2)
	-0.47	Psychomotor Vigilance Task	Probe	71	30	(Unsworth & Robison, 2018, Exp. 3)
	-0.32	Psychomotor Vigilance Task	Probe	81	23	(Unsworth & Robison, 2018, Exp. 4)
	-0.2	Whole-report Task	Probe	36	8	(Robison & Unsworth, 2019, Exp. 2)
	-0.07	WM Whole-report Task	Probe	38	18	(Robison & Unsworth, 2019, Exp. 3)

Task Evoked Pupil Response (Change Rate)	-0.87 ^a	Stop-signal	Probe	20	40	(Mittner et al., 2014)
--	--------------------	-------------	-------	----	----	------------------------

Notes.

1. This table does not list all local measures examined in the previous literature. A measure was listed here either because (1) its Cohen's d_z absolute value was greater than .20 or (2) more than one study examined this measure. These criteria excluded many measures that were examined in a single study or had a low effect size.
2. Cohen's d_z : Positive values indicate MW > On-task; Negative values indicate MW < On-task. Values were obtained from the original papers when possible. Superscript ^a indicates that the value was calculated from test statistics (e.g., t , F , and $Chi-squared$); superscript ^b indicates that the value was calculated from means and standard deviations provided in the paper (assuming within-subject correlation is .5).
3. Method: "Probe" indicates studies that used the probe-caught method; "Self" indicates studies that used the self-caught method.

1.5 Summary and Dissertation Outline

MW is ubiquitous and is often associated with reduced task performance. However, due to its covert nature, MW often lacks obvious external indications, hindering further understanding of this phenomenon. Recent studies show that MW, cognition decouples from the external input, and this attentional decoupling process is associated with changes in gaze behaviors. However, most of the studies used reading comprehension as their primary task, and there is quite a bit of inconsistency regarding how MW was defined among different studies. To further understand the relationships between task performance, eye movements, and MW, the current dissertation presents three studies with distinct task settings: visual search, scene perception, and reading comprehension. Here, I outline the aims of each study:

Study 1 examined how MW might affect visual search performance, and particularly the ability to suppress salient distractors during visual search. A unique feature of this study is to use pupil size as an objective indication of attentional states, in addition to subjective reports. This study is interested in whether pupil size and self-reported MW are correlated with each other, and whether they are associated with the same type of performance deficit.

Study 2 used a scene encoding task to study how MW might affect the way eye movements change over time and their relationship with scene content. A unique feature of this study is to explore the sequential aspects of eye movements (i.e. scanpaths) during MW.

Study 3 examined how MW might affect the reader's ability to detect semantic incongruities in the text and make necessary revisions of their understanding. Different from many previous studies that used reading tasks, this study used "garden-path jokes" (Dyner, 2009) as reading material. Garden-path jokes provide an interesting opportunity to study how online language processing is affected by different attentional states.

Throughout the tasks, the definitions of MW were held constant. Therefore, we can also examine if there are underlying similarities in the process of MW across different task scenarios. The final chapter of this dissertation explored if there are context-independent (i.e., global) eye movement features of MW.

Taken together, the dissertation may expand our understanding of MW to a wider range of tasks and examine the feasibility of using task-general eye movement features to detect episodes of MW.

Chapter 2 Mind-wandering during Visual Search

2.1 Abstract

Our attention wanders from the task at hand from time to time, and these fluctuations of attention may affect our task performance. We examined if off-task attentional states are associated with performance deficits in visual search, particularly an increased interference from salient distractors. Participants searched for a target shape among heterogeneous shapes while a visually salient distractor (a non-target shape in a distinct color) appeared on 50% of the trials. Thought probes and pre-trial pupil size consisted of subjective and objective measures of participants' attentional states, respectively. Overall, participants could effectively suppress early eye movements to the distractor, and this oculomotor suppression effect did not vary as a function of self-reported attentional states. Participants were also equally fast in their first look to the target during mind-wandering and on-task episodes. But the presence of the distractor increased time looking at the target during mind-wandering compared to when on-task, and this processing cost contributed to increased response time in distractor-present trials during mind-wandering. Moreover, there was a U-shaped relationship between pre-trial pupil size and response time for both distractor-present and distractor-absent trials, such that faster response times were achieved with intermediate pupil sizes. However, the elevated response times at extreme levels of pupil size cannot be attributed to increased looking time on the target (as we found for the self-reported mind-wandering trials). Pre-trial pupil size was also not associated with self-reports of mind-wandering. In sum, there was evidence for increased distractor interference during self-reported mind-wandering, although the interference was not apparent

during initial eye movements. Pre-trial pupil size can to some extent predict task performance in the upcoming trial, but it may measure different processes of task disengagement from those measured by thought probes.

2.2 Introduction

Successful task performance often requires us to attend to task-relevant information while ignoring irrelevant information. However, our attention can be easily captured by salient-but-irrelevant objects in the environment. For example, when asked to search for an item with a unique shape (e.g., a diamond among circles, or *vice versa*), participants had slower response times if one of the non-target items had a unique color (Theeuwes, 1992). Participants' eye movements are also captured by the salient distractor as measured by initial eye movements (Gaspelin, Leonard, & Luck, 2017, exp. 1; Theeuwes, 2010). Participants knew that the color dimension was irrelevant to the task but were somehow still distracted by it. This distractor presence cost is a robust indication of perceptual distraction in visual search.

If salient distractors always involuntarily capture attention, people will have difficulties focusing on the task at hand. Fortunately, individuals' top-down modulation of attention also plays a key role in determining whether distractors capture attention (Bacon & Egeth, 1994; Folk & Remington, 1998; Lamy & Tsal, 1999; Leber & Egeth, 2006). The *signal suppression hypothesis* provides a theoretical attempt to integrate top-down and bottom-up influences on attention, highlighting the important role of inhibition in avoiding salient distractors (for a review, see Gaspelin & Luck, 2018b). According to this theory, salient stimuli have the intrinsic ability to capture attention, but capture can be avoided if an individual exerts top-down inhibition. The theory further claims that salient distractors can be suppressed below the baseline level (relative to non-salient distractors) when inhibition is exerted. In the original version of the

color singleton task (Theeuwes, 1992), avoiding distraction can be difficult due to the task's color and shape configuration (Bacon & Egeth, 1994; Leber & Egeth, 2006). But in a variation of the task that promotes top-down control, behavioral and eye-tracking data converged to show evidence of distractor inhibition, which can eliminate or even reverse the singleton presence cost (Gaspelin, Leonard, & Luck, 2015; Gaspelin et al., 2017; Gaspelin & Luck, 2018a). For example, Gaspelin et al. (2017) observed a distractor presence *benefit* (i.e., faster response times when the color singleton was present vs. absent) in response times in their revised color singleton paradigm. They also reported that initial saccades were *less* likely to land on the distractor compared to non-salient distractors, even for the fastest saccades which are known to be sensitive to visual saliency (Anderson, Ort, Kruijne, Meeter, & Donk, 2015). The distractor presence benefit, according to the authors, indicates that the distractor position was inhibited from the visual search, effectively decreasing the number of items to be searched (Gaspelin & Luck, 2018b).

While empirical evidence clearly shows that inhibiting salient distractors is possible, it is important to consider the fact that our ability to regulate attention waxes and wanes, which may result in fluctuations in inhibitory performance. Gaspelin and Luck (2018, p.83) explicitly stated that one of the premises for distractor inhibition is that “participants are in a state of good attentional control.” However, our ability to maintain control over our attention and focus fluctuates from time to time (deBettencourt, Keene, Awh, & Vogel, 2019; Kam et al., 2011; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997; Smallwood et al., 2004). In particular, our mind often wanders away from the ongoing task to internal and task-unrelated thoughts, a phenomenon called *mind-wandering* (MW; Smallwood & Schooler, 2006). Studies have shown that MW occurs frequently (30% ~ 50%) of the time in both lab and real-life settings (Kane et

al., 2007; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012; Seli, Cheyne, et al., 2015; Varao-Sousa & Kingstone, 2019). MW occurs when executive-control processes fail to suppress the interference created by task-unrelated thoughts (Kane & McVay, 2012; McVay & Kane, 2009) and often leads to lower performance in attention-demanding tasks (Mooneyham & Schooler, 2013).

One important reason for the lowered task performance during MW is that the top-down control of attention that is often critical to good task performance is impaired (Smallwood & Schooler, 2006). For example, Reichle, Reineberg, and Schooler (2010) showed that during MW, the usual association between word frequency and fixation duration was absent, suggesting that top-down cognition was not actively controlling the eyes to extract lexical information from the words. Similarly, Mills, Graesser, Risko, and D’Mello (2017) showed that during MW, readers were less able to adjust their reading speed based on text difficulty, signaling a loss of top-down control of comprehension. Seli, Cheyne, and Smilek (2013) asked participants to respond to a continuous rhythmic presentation of tones via button presses and found that during MW, participants’ responses desynchronized with the tones. A handful of studies have also shown that MW was associated with poor response inhibition and increased response variability in sustained attention tasks (Bastian & Sackur, 2013; Cheyne, Solman, Carriere, & Smilek, 2009; Christoff et al., 2009; McVay & Kane, 2012; Smallwood et al., 2004; Stawarczyk, Majerus, Catale, & D’Argembeau, 2014). These findings collectively show that MW is associated with deficits in the control of attention and behavioral responses.

Because MW is a state of low-level top-down control, visual search performance, in particular the ability to suppress a salient distractor, might be impaired during MW. The signal suppression hypothesis predicts that attention and eye movements would be driven by visual

saliency if no top-down inhibition is imposed (Gaspelin & Luck, 2018b). Therefore, one might expect the distractor to capture attention and eye movements during MW due to a lack of top-down control. Interestingly, some previous studies show that high levels of MW were associated with reduced processing of both the target and the distractor. For example, in a continuous performance task, Kam et al. (2011) found that MW was associated with reduced cortical processing of the target and the occasional onset of a novel stimulus, as indicated by reduced event-related potentials (the P3 component for the target and the P1 component for the novel stimulus). Similarly, Barron et al. (2011) found that individuals with high levels of MW in a visual oddball task showed reduced cortical processing of both the target and the distractor. Esterman et al. (2013) also showed reduced distractor processing during periods of lapses of attention compared to focused periods. These findings, according to Schooler et al. (2011), suggest that a global reduction in perceptual processing occurs during MW, regardless of the task-relatedness of the stimuli. Therefore, it is also possible that, during MW, the salient distractor captures no more attention compared to a non-salient item that is not the target. In sum, we hypothesize that the distractor suppression effect should disappear during MW, but it is unclear whether the distractor would additionally create a capture effect.

2.2.1 Pupil size as an objective measure of task focus.

MW has been typically measured via the thought probe technique: the task occasionally pauses to ask participants to self-report their current attentional state. Recent studies show the promise of using pupil size as an objective measure of attentional state (e.g., Franklin, 2013; Grandchamp et al., 2014; Konishi et al., 2017; Unsworth & Robison, 2016, 2018). Pupil size has been traditionally used as a measure of cognitive effort (Kahneman & Beatty, 1966) and is recently linked to activation in the locus coeruleus (Aston-Jones & Cohen, 2005; Gilzenrat et al.,

2003, 2010; Murphy et al., 2011; Rajkowski et al., 1993). The locus coeruleus (LC) is responsible for the synthesis of norepinephrine and plays an important role in regulating arousal, attention, and responses to stress (Benarroch, 2009). A large body of research has reported an inverted-U relationship between tonic LC activity and task performance, such that optimal performance is achieved with intermediate levels of tonic LC activity; low tonic LC activity is associated with drowsiness, whereas high LC activity is associated with task disengagement (Aston-Jones & Cohen, 2005; Berridge & Waterhouse, 2003). Critically, empirical studies have found a close relationship between tonic LC activity and tonic pupil diameter, although the relationship is not yet fully understood (Aston-Jones & Cohen, 2005; Rajkowski et al., 1993). Relatedly, several studies have found that poor task performance is associated with both relatively small and relatively large pupil sizes, suggesting an inverted-U relationship (Gilzenrat et al., 2010; Kristjansson et al., 2009; Murphy et al., 2011; Smallwood et al., 2012; van den Brink et al., 2016). These findings suggest that pupil size can be used as an objective measure of attentional states in the current task, in addition to the thought probes. It is worth noting that the relationship between subjective reports of MW and tonic pupil size are highly inconsistent in the previous literature. MW was reported to be associated with relatively small (Grandchamp et al., 2014; Konishi et al., 2017; Mittner et al., 2014; Unsworth & Robison, 2016, 2018) or relatively large pre-trial pupil size (Jubera-García et al., 2019, exp.1; Smallwood et al., 2011, 2012), but sometimes MW was not associated with changes in pre-trial pupil size at all (Jubera-García et al., 2019, exp. 2; Robison & Unsworth, 2019; Unsworth & Robison, 2018, exp. 2). These results are broadly consistent with the idea that MW is a heterogenous term (Seli, Kane, Metzinger, et al., 2018) and may emerge under different levels of arousal (Lenartowicz, Simpson, & Cohen, 2013).

2.2.2 The current study.

The current study has two general purposes: (1) to examine if MW is associated with reduced visual search performance, and in particular, reduced inhibition of salient distractors, and (2) to examine if pre-trial pupil size can predict visual search performance and subjective reports of MW. The visual search task closely followed Experiment 2 in Gaspelin et al. (2017). Participants were asked to find the target shape (either a diamond or a circle; counterbalanced across participants) among heterogeneous shapes. On 50% of the trials, one of the non-target shapes was drawn in a different color (a red item among green ones or a green item among red ones; counterbalanced across participants), thus becoming the salient distractor. Participants' task was to respond to the orientation of the line segment inside the target (either horizontal or vertical). Eye movements and pupil size were collected as participants completed the task. Thought probes appeared after some of the trials to obtain subjective reports of MW.

As noted previously, MW is a heterogeneous phenomenon that encompasses many different forms. For example, previous studies often differentiate between intentional mind-wandering and unintentional MW (Seli, Risko, Smilek, et al., 2016). Intentional MW is a controlled and deliberate process of disengaging from the task. On the other hand, unintentional MW refers to task disengagement that is out of an individual's control. Some studies have shown that both types of MW are associated with lower task performance (Seli, Cheyne, et al., 2015; Seli, Wammes, Risko, & Smilek, 2016). In other cases, intentional and unintentional MW are dissociable (e.g., Golchert et al., 2017; Phillips et al., 2016; Seli, Risko, & Smilek, 2016; Seli, Smallwood, et al., 2015). For example, ADHD, who is known to lead to heightened distractor interference (Forster, Robertson, Jennings, Asherson, & Lavie, 2013; Mason, Humphreys, & Kent, 2005), is associated with increased unintentional MW but not intentional MW (Seli,

Smallwood, et al., 2015). Currently, there is not enough information to make specific predictions about differences in distractor inhibition for different types of MW. Moreover, given that intentional MW may have a lower occurrence rate than unintentional mind-wandering (Seli, Risko, & Smilek, 2016), the statistical power of tests involving intentional MW might be limited. Despite these challenges, it may still be worthwhile to distinguish between intentional and unintentional MW, and both were measured in the current study for exploratory purposes.

We pre-registered our study on AsPredicted which can be found at <http://aspredicted.org/blind.php?x=w7cz8p>. Unless otherwise noted, all methods and analyses proceeded as in the pre-registered plan.

2.3 Methods

2.3.1 Participants.

We determined the sample size based on a power analysis using the *anova_power* package (http://shiny.ieis.tue.nl/anova_exact/; Lakens & Caldwell, 2019). Gaspelin et al. (2017, Experiments 2-3) reported an oculomotor suppression effect of $d_z = 1.63$. While there might be a reversed effect (oculomotor capture) during MW, a parsimonious estimation would be no effect ($d_z = 0$). Moreover, because the analyses will be based on a subset of trials (probed trials), the actual effect size might be lower compared to using full trials. Thus, we expect a suppression effect with $d_z = .8$ for on-task condition, and a $d_z = 0$ for the off-task condition. With alpha = .05 and a hypothesized correlation of .3 among within-subject factors, simulation results show that a sample size of 40 can achieve over 80% power for both the interaction between attention and distractor presence as well as the post-hoc tests. The planned sample size is also comparable to previous studies examining the relationship between pupil size, task performance, and MW (e.g.,

Franklin, 2013; Jubera-García et al., 2019; Konishi et al., 2017; Robison & Unsworth, 2019; Smallwood et al., 2011; Unsworth & Robison, 2016, 2018).

We recruited 42 undergraduates (Mean Age: 18.76, SD Age: .99, 62% female) from the University of Michigan to participate in this study for course credit. We excluded participants based on the criteria listed in the pre-registered plan, with one exception: in the protocol, we stated that participants should be excluded if they “have consumed nicotine, alcohol, and other drugs that may affect their attention and memory within the past 12 hours”. It turns out that a substantial portion of our participants had consumed nicotine ($n = 3$), alcohol ($n = 2$), or caffeine ($n = 19$) in the past 12 hours. Given our constraints in time and personnel, this exclusion criterion was not enforced. Six participants were excluded due to incompleteness and three were excluded due to pre-existing medical conditions. The final sample size consisted of 33 participants.

2.3.2 Apparatus.

The visual search task was presented on a 20.1-inch computer screen at approximately 80 centimeters to the participant. Binocular eye movements were recorded by an EyeLink 1000 System at a sampling rate of 500 Hz. A chin rest was used to ensure tracking stability. The experiment was implemented using the OpenSesame software (Mathôt, Schreijf, & Theeuwes, 2012) with functions from the PyGaze package (Dalmaijer, Mathôt, & der Stigchel, 2014).

2.3.3 Tasks and stimuli

2.3.3.1 Visual search task.

The visual search task was programmed based on Gaseplin et al.’s (2017) experiment 2. The search array consisted of 6 items distributed equally on an imaginary circle with a radius of 4.5° (see Figure 2-1). The target was always a diamond ($0.8^\circ \times 0.8^\circ$) for half of the participants

and a circle (0.9° radius) for the other half. The target location was randomized across trials. Non-target items consisted of squares, hexagons, and the unselected target item (circle for the diamond group, diamond for the circle group). These non-target items were selected randomly for each trial with the constraint that each shape did not appear more than twice in the search array. For half of the trials, all items in the search array appeared in either red or green, counterbalanced across participants. For the other half of the trials, one randomly selected non-target item appeared in the opposite color, thus becoming the color singleton. Inside each item, there will be a line segment ($.37^\circ$) that was drawn in black and had an orientation of either vertical or horizontal (randomly chosen). The task was to identify the orientation of the line segment inside the target item. Pilot testing showed that the task's display setting encourages participants to make an eye movement to the target to identify the line segment. Participants were asked to ignore the possible color singleton because "it will never be the target".

A trial started with a fixation cross at the center of the screen. Participants must maintain their gaze within a 1.5° radius of the fixation cross for 2000 ms to trigger the search array. Then, the fixation cross disappeared, and the search array appeared and remained visible until response. If the fixation cross failed to trigger within 10 seconds, the search array would automatically appear. Participants responded to the orientation of the line inside the target item using the "z" key and the "/" key. After a response, the screen remained dark for about 600 ms until a new trial started. Participants first completed a practice block of 80 trials, followed by 6 experimental blocks of 80 trials each. Thought probes (discussed more below) followed 10% percent of the trials in each block.

The eye-tracker was calibrated at the beginning of each block. Participants were asked to focus on a white dot appearing at the top, bottom, left, right, and the center of the screen while

the experimenter manually confirmed the fixated locations by pressing the spacebar on the host computer. Then, a validation procedure was performed to evaluate the spatial errors of the calibration. The eye-tracker was re-calibrated and re-validated after every trial that had more than 10 seconds of fixation-triggering time.

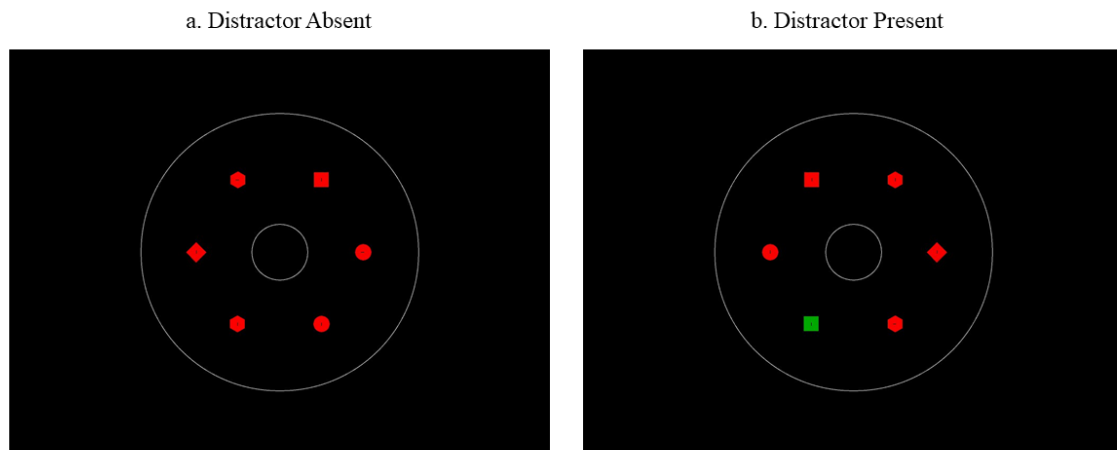


Figure 2-1. Example trials with and without the color singleton. In this figure, the target is the diamond. On 50% of the trials, one of the non-target items will appear in a different color (see right panel). The shape of the color singleton may vary from trial to trial, but it will never be the target shape. Participants' task is to indicate the line orientation inside the target item. Pilot studies suggested that the study's display setting encourages participants to make an eye movement to the target to identify the line segment. Previous studies using a similar configuration (Gaspelin et al., 2017) showed that participants were able to suppress initial eye movements to the color singleton. The two solid white circles represent the boundaries of the area of interest and were not displayed in the actual task.

2.3.3.2 Thought probes.

A thought probe occurred after 10% of the trials at pseudo-random locations in each block. Distractor-present and distractor-absent trials were probed equally often, and each probed trial was preceded by at least 3 and at most 20 non-probed trials.

The probe asked participants to “select the one that best describes your conscious experience during the previous trial”. Participants could choose either “on-task” by pressing the “up arrow” key, “unintentional mind-wandering” by pressing the “left arrow” key, or “intentional mind-wandering” by pressing the “right arrow” key. Intentional mind-wandering

was defined as "*you intentionally decided to think about things that are unrelated to the task*", while unintentional mind-wandering was defined as "*your thoughts drifted away despite your best intentions to focus on the task*" (Seli et al., 2015). After the probes, another screen showed up to prompt the participant to move their hands back to the response keys. Participants pressed the spacebar to resume the task.

2.3.3.3 To-do list activity.

This activity was adopted from Kopp, D'Mello, and Mills (2015) and was intended to increase the level of MW in a subsequent task. Participants were given 5 minutes to make a list of things they plan to do in the next five days. All participants completed this task before the visual search task.

2.3.3.4 Questionnaires.

Participants completed several questionnaires after the visual search task: the Conners' Adult ADHD Rating Scale (Conners, Erhardt, & Sparrow, 1999), the Cognitive Failures Questionnaire (Broadbent, Cooper, FitzGerald, & Parkes, 1982), and the White Bear Suppression Inventory (Wegner & Zanakos, 1994). Data from these questionnaires were not analyzed in this study.

2.3.4 Procedure.

After signing the consent form, participants first completed a health and demographics questionnaire, in which we collected their basic demographic information and medical history. Then, we asked participants to complete the Ishihara color deficiency test (Ishihara, 2010) to screen for color-blindness. After that, participants completed the to-do list activity. Immediately

after, the visual search task started. Participants took a short break in between the blocks. After the visual search task, participants completed the three questionnaires described above. Finally, participants were debriefed and thanked.

2.3.5 Data Analysis.

2.3.5.1 Pre-processing of pupil data.

Pupil size was recorded as the surface area in scaled image pixels, per Eyelink's default. Pupil size data recorded during the fixation cross of each trial went through several pre-processing steps. First, we removed blinks using a novel algorithm developed by Hershman et al. (2018). Second, we removed pupil outliers based on location and size. Specifically, pupils with locations away from the central fixation cross (defined as pupils outside of a $3^{\circ} \times 3^{\circ}$ region at the center of the screen) were removed; pupil sizes outside of 2.5 standard deviations from the mean were also removed (note that if the standard deviation was below $0.1 \times \text{mean}$, this procedure was not performed to prevent erroneous removal of data in a very steady signal). Third, we performed a Hanning smoothing of the pupil signal with a window size of 10 samples. Fourth, we performed a linear interpolation of the pupil signal to fill in missing data. Finally, we down-sampled the pupil signal to 10 Hz. The code for pupil pre-processing is available at <https://github.com/HanZhang-psych/Pupil-preprocessing>.

2.3.5.2 Pre-processing of fixations.

Fixations that were greater than 1500 ms or shorter than 80 ms were removed from analysis. To determine the landing position of fixations during the search array, we defined an annulus-shaped area of interest (see Figure 2-1) with an inner radius of 1.5° and an outer radius of 7.5° . Fixations inside the annulus were assigned to the closest item. We defined the initial

landing position as the location of the first fixation that landed on any of the search items. As stated in the pre-registered protocol, we also defined two supplemental measures: the total looking time on each item (fixation sum on the target, distractor, and the non-targets) and the time needed to locate the target (the latency from the trial start to the first fixation on the target). These measures may be informative about what happens after the initial capture/suppression because they include later eye movements.

2.3.5.3 Exclusion of trials.

As stated in the pre-registered protocol, trials were excluded (1) if they had over 30% of tracking loss (2.7% of trials), (2) if they had a response time over 3 standard deviations from the mean (1.6% of trials), (3) if they had a pre-trial pupil size over 3 standard deviations from the mean (0.4% of trials). No additional participants were excluded due to a low number of valid trials or low response accuracy. The overall accuracy was quite high (97%) and did not significantly differ in any of our analyses. Therefore, incorrect trials were also discarded.

2.3.5.4 Statistical Analysis.

We combined intentional and unintentional MW as a single category in statistical testing, because we performed a power analysis based on the combined category, and splitting MW into two categories may impact the statistical power of our tests. Indeed, the rate of intentional MW was only 10% in our study (the rate of unintentional MW was 30%). Therefore, any statistical tests involving intentional MW may suffer from increased measurement error. For exploratory purposes, we simply plotted the data for intentional and unintentional MW separately without performing any significance testing. Statistical analyses were performed using linear mixed models (Bates, Mächler, Bolker, & Walker, 2014). Unless otherwise noted, all models included

all possible random intercepts and slopes justified by the design (Barr, Levy, Scheepers, & Tily, 2013).

2.4 Results

2.4.1 Distractor suppression across all trials.

We first examined if we could replicate the distractor suppression effect observed in Gaspelin et al. (2017) across all trials. Note that this is not a pre-registered analysis, but the study was sufficiently powered to test this effect. We found that distractor-present trials (*Mean* = 1034.54, *SD* = 217.57) on average had a faster response time compared to distractor-absent trials (*Mean* = 1047.40, *SD* = 217.55), and the difference was marginally significant, $t(32) = 1.94$, $p = .06$, $d_z = .34$. Initial fixations during distractor-present trials were less likely to land on the distractor (Mean = 9.07%, SD = 4.96%) compared to the rest of the non-target items averaged (Mean = 12.80%, SD = 2.84%), $t(32) = -4.24$, $p < .001$, $d_z = .74$. Therefore, we replicated the key findings in Gaspelin et al. (2017) using all of our trials.

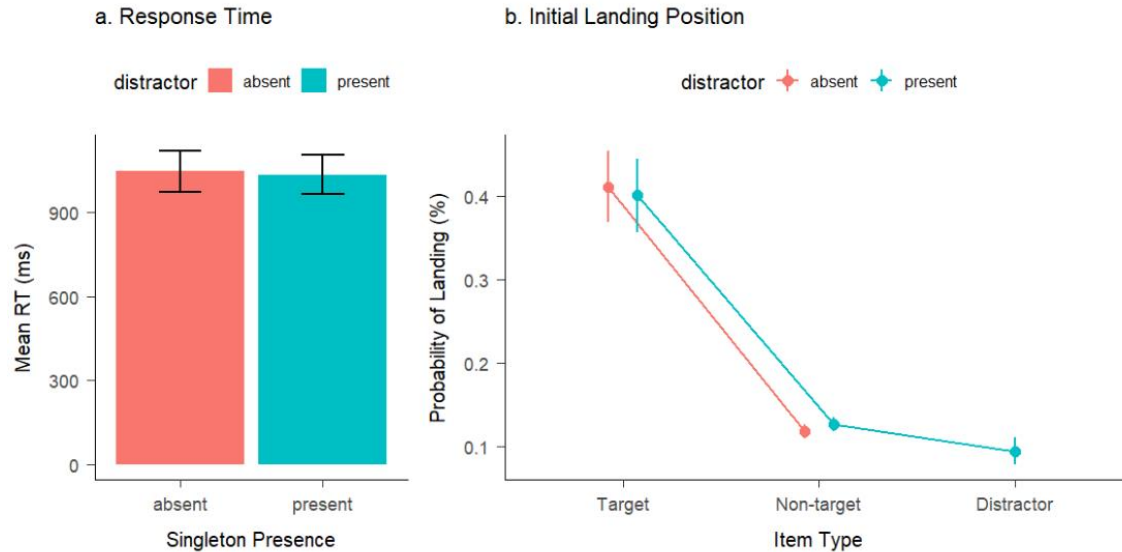


Figure 2-2. Distractor suppression effect across all trials, as illustrated by response time (panel *a*) and initial landing position (panel *b*). Panel *a* shows that distractor-present trials were on average faster (marginally significant) compared to distractor-absent trials. Panel *b* shows that participants were less likely to have their first eye movements on the distractor than on an average non-target item in distractor-present trials. Error bars showed 95% confidence intervals.

2.4.2 MW over time.

Overall, 59.6% of the probed trials were classified as being on-task, 30.2% were classified as unintentional MW, and 10.1% were classified as intentional MW. As stated in the pre-registration, we explored how MW changed over time as a function of block number. We conducted a growth curve analysis using orthogonal polynomials to model the linear and quadratic trend of on-task reports. We found a marginally significant linear trend, $b = -.06$, $t = -1.89$, $p = .07$, and a significant quadratic trend, $b = .12$, $t = 3.87$, $p < .001$, indicating a rise in on-task reports towards the end of the blocks.

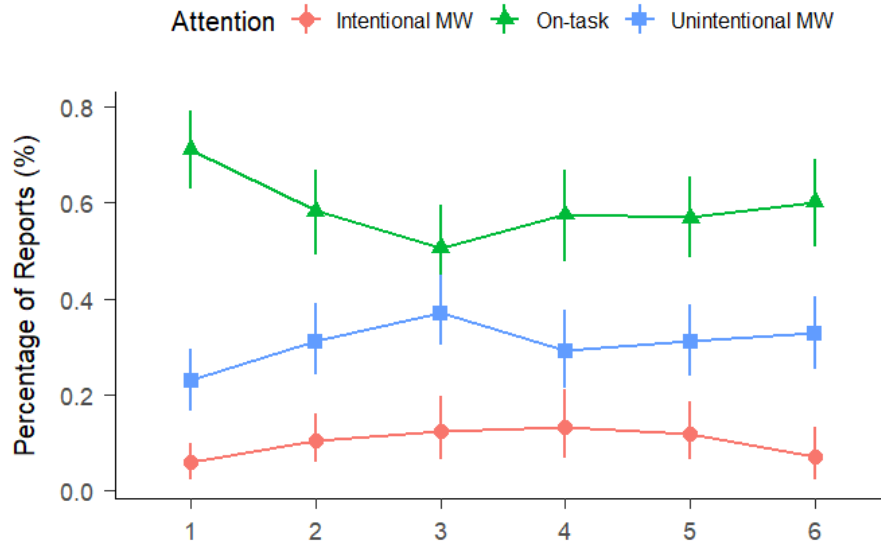


Figure 2-3. Subjective reports of on-task and MW throughout the 6 experimental blocks. Error bars showed 95% confidence intervals. There were a marginally significant linear trend and a significant quadratic trend for the time course of on-task reports.

2.4.3 Visual search performance during MW.

2.4.3.1 Response time.

We examined distractor interference among the probed trials as a function of self-reported attentional states. A two-by-two repeated measure ANOVA with attention (on-task vs. MW), distractor presence (present vs. absent) reveals a significant main effect of attention, $F(1, 94.88) = 5.22, p = .02$, such that response times were longer during MW than during on-task trials (marginal difference = 76.5 ms). The main effect of distractor was not significant, $F(1, 94.19) = .03, p = .87$. There was a marginally significant interaction effect, $F(1, 94.19) = 3.15, p = .08$. Post-hoc analysis showed that response time in distractor-absent trials did not significantly differ between on-task and MW, $b = 17.4, t = .37, p = .71$, but there was a significant difference for distractor-present trials, $b = 135.6, t = 2.87, p = .006$. Therefore, when the distractor was present,

participants responded more slowly during MW compared to when on-task. Data with intentional and unintentional MW plotted as separate categories are shown in Figure 2-4, panel *a*.

2.4.3.2 Initial landing position.

The increased response time on distractor-present trials during MW might be caused by an oculomotor capture effect by the salient distractor. To test this possibility, we conducted a two-by-two repeated measure ANOVA with attention (on-task vs. MW) and item type (distractor vs. non-target). Results showed a significant main effect of item type, $F(1, 32.27) = 11.29, p = .002$, indicating an overall oculomotor suppression effect (marginal difference = 4.8%). The main effect of attention was not significant, $F(1, 32.04) = .01, p = .92$. The interaction term was also not significant, $F(1, 32.21) = .56, p = .46$. Therefore, there was no evidence of oculomotor capture by the salient distractor during MW. Data with intentional and unintentional MW as separate categories are shown in Figure 2-4, panel *b*.

2.4.3.3 Time to target.

Another possibility for the increased response time during MW might be that participants were generally slower to locate the target. However, a two-by-two repeated measure ANOVA with attention (on-task vs. MW) and distractor presence (present vs. absent) showed no significant main effects (attention: $F(1, 63.18) = .31, p = .58$; distractor presence: $F(1, 32.21) = 1.22, p = .28$) and no significant interaction ($F(1, 63.53) = 2.38, p = .13$). Data with intentional and unintentional MW as separate categories are shown in Figure 2-4, panel *c*.

2.4.3.4 Total looking time.

A third possibility for the increased response time is a potential delay after the target has been located, which can be captured by the total looking time on the target. We analyzed total looking time on the distractor, the target, and the rest of the items (non-targets) with separate ANOVAs. For total looking time on the distractor, there was no significant difference between on-task and MW trials, $F(1, 17.43) = .29, p = .59$. For total looking time on the target, a two-by-two repeated measure ANOVA with attention (on-task vs. MW) and distractor presence (present vs. absent) showed that significant main effect of attention, $F(1, 94.66) = 12.33, p < .001$, such that total looking time was overall longer during MW than during on-task episodes (marginal difference: 50.7 ms). There was also a significant main effect of distractor presence, $F(1, 94.18) = 5.04, p = .03$, such that the total looking time on the target was overall shorter when the distractor was absent than when it was present (marginal difference: -32.2 ms). Importantly, this effect was qualified by a significant interaction, $F(1, 94.18) = 9.70, p = .003$. Post-hoc analysis showed that, when the distractor was present, participants spent significantly longer time looking at the target during MW compared to when on-task, $b = 95.38, t = 4.68, p < .001$, but a similar pattern was not found for distractor-absent trials, $b = 6.01, t = .30, p = .77$. Finally, for total looking time on non-targets, there was only a marginally significant effect of distractor presence, $F(1, 32.71) = 3.27, p = .08$, which is likely to reflect the fact that there was one less of such items in the distractor-present trials (because one of the non-targets becomes the color singleton). The main effect of attention was not significant, $F(1, 29.91) = .36, p = .56$. The interaction between attention and distractor presence was also not significant, $F(1, 31.46) = .65, p = .43$. Data with intentional and unintentional MW as separate categories are shown in Figure 2-4, panel *d*.

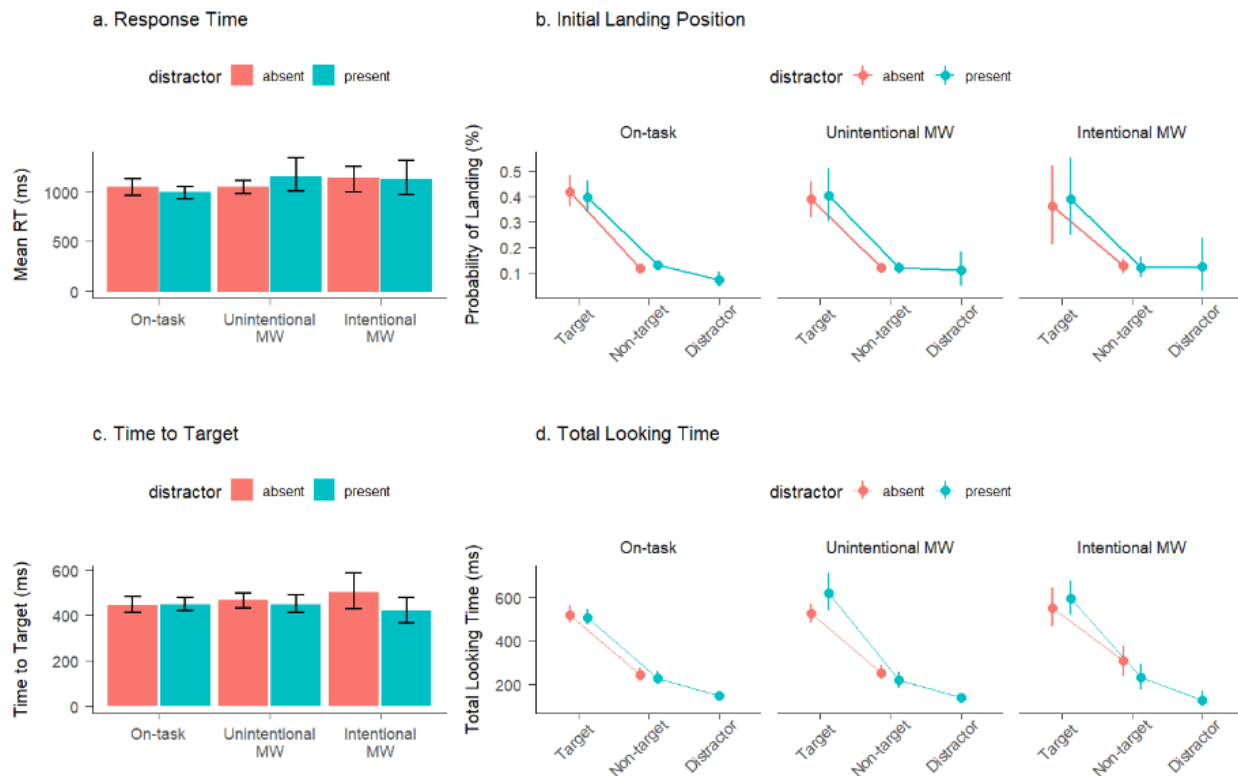


Figure 2-4. Response time and eye movement measures for trials reported as on-task, unintentional MW, and intentional MW. Panel *a* shows response time. Response times were overall longer during MW (intentional and unintentional combined) than during on-task episodes. Panel *b* shows the initial landing position. There was a main effect of distractor suppression (as measured by the difference between non-target and distractor for distractor-present trials), but the effect did not vary between on-task and MW trials. Panel *c* shows the time spent to locate the target. There were no significant main effects or a significant interaction. Panel *d* shows the total looking time. Participants during MW spent longer time looking at the target if a distractor was present compared to when it was absent. But this was not the case when participants were on-task. Overall, these results suggest that participants during MW could still suppress initial eye movements to the salient distractor and they seemed to be equally as fast at locating the target compared to when they were on-task. However, the presence of the distractor somehow slowed down their processing of the target, causing longer looking times and possibly slower response times.

2.4.4 Pupil size over time.

So far, we have been focusing on subjective reports of attentional states. Now we turn to pre-trial pupil size, an objective measure of participants' attention at each trial. As stated in the pre-registration, we first explored how pupil size (z-scored within each participant) changed as a function of block number. We conducted a growth curve analysis with orthogonal polynomials. We found a significant linear trend, $b = -.61$, $t = -4.33$, $p < .001$, and a significant quadratic trend, $b = .26$, $t = 2.45$, $p = .02$. Therefore, while pupil size decreased throughout the experiment, the

change did not follow a linear trend (but note that pupil size was recorded as the surface area, not diameter). The data are plotted in Figure 2-5.

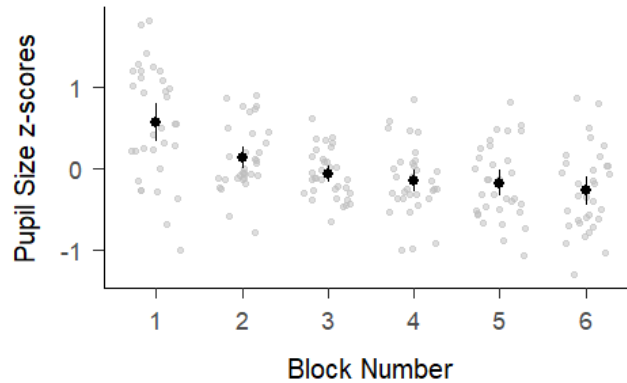


Figure 2-5. Pupil size (z-scored within each participant) as a function of block number. Error bars show 95% confidence intervals.

2.4.5 Visual search performance and pupil size.

2.4.5.1 Response time.

As discussed in the introduction, previous literature suggests an inverted-U relationship between pupil size and task performance. Therefore, we tested whether relatively slow response times were associated with both relatively small and relatively large pupil sizes. Specifically, we divided each participant's pupil sizes into ten deciles and plotted them against each participant's response time z-scores (z-scores were used to reduce individual differences). The plot is shown in Figure 2-6, panel *a*. The smoothed curve appears to show a quadratic pattern. We formally tested this pattern using a growth curve analysis. Orthogonal polynomials (linear and quadratic), distractor presence (present vs. absent [baseline]) and their interaction entered the model as fixed effects. The linear trend was not significant, $b = .07$, $t = .14$, $p = .89$. However, there was a significant quadratic trend, $b = 1.03$, $t = 3.48$, $p = .001$. Moreover, there was a significant effect of distractor presence, $b = -.05$, $t = -3.71$, $p < .001$, which indicates an overall distractor presence benefit. Distractor presence did not significantly interact with either the linear term ($b = .36$, $t =$

.91, $p = .37$), or the quadratic term ($b = -.21$, $t = -.57$, $p = .57$), suggesting that distractor-present and distractor-absent trials had similar relationships with pupil size.

2.4.5.2 Initial landing position.

We examined if there was also increased oculomotor capture by the distractor at extreme levels of pupil size. For each pupil size decile, we calculated the percentage of trials with the initial landing position on the distractor. The data are shown in Figure 2-6, panel *b*. A growth curve analysis showed that the linear trend was not significant, $b = .04$, $t = .65$, $p = .52$, but there was a significant quadratic trend, $b = .13$, $t = 2.06$, $p = .04$.

2.4.5.3 Time to target.

Similar to response times, we plotted the z -scores of time to target as a function of pupil size deciles (see Figure 2-6, panel *c*). Growth curve analysis showed that the linear trend was not significant, $b = -.06$, $t = -.12$, $p = .91$. There was, however, a significant quadratic trend, $b = .98$, $t = 2.56$, $p = .02$. All other effects were not significant, $ps > .11$. Therefore, there was also a quadratic relationship between pupil size and the time needed to locate the target, and this relationship did not significantly differ between distractor-present and distractor-absent trials.

2.4.5.4 Total looking time (on the target).

Our analyses of self-reported MW showed a distractor interference on the time looking at the target item. Can the same distractor interference explain the elevated response times at relatively small and relatively large pupil sizes? To answer this question, we plotted the z -scores of the total looking times on the target as a function of pupil size deciles (Figure 2-6, panel *d*). A growth curve analysis found that the linear trend was not significant, $b = .56$, $t = .96$, $p = .34$. The

quadratic term was also not significant, $b = .61$, $t = 1.49$, $p = .15$. There was also no global difference between distractor present vs. absent, $b = -.02$, $t = -1.17$, $p = .24$. There was a marginally significant interaction between distractor presence and the linear trend, $b = .82$, $t = 1.75$, $p = .09$, suggesting a stronger linear trend for distractor-present trials compared to distractor-absent trials. However, the interaction between distractor presence and the quadratic trend was not significant, $b = -.51$, $t = -.92$, $p = .36$. Therefore, total looking time on the target did not follow a quadratic relationship with pupil size for both distractor-present and distractor-absent trials. The elevated response times at extreme levels of pupil size cannot be attributed to participants looking at the target for a longer duration.

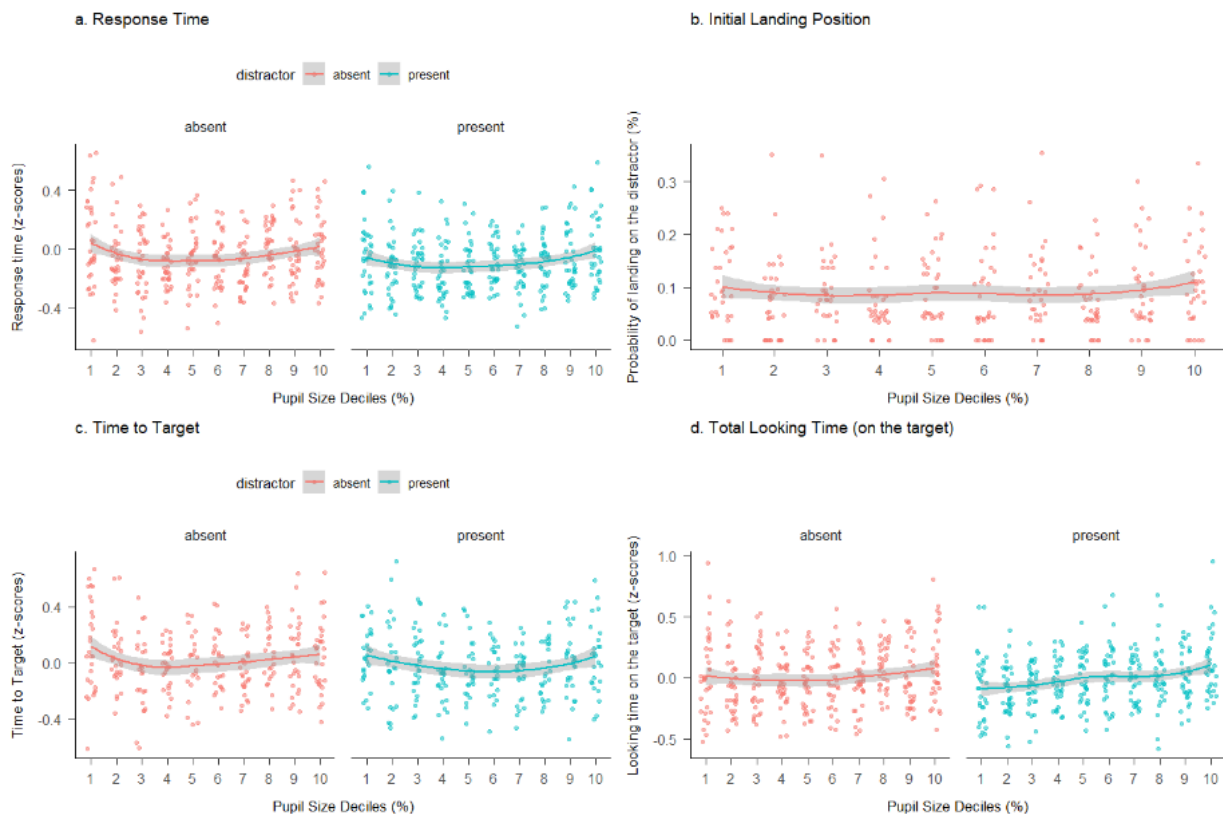


Figure 2-6. The relationship between pupil size and several performance measures. Pupil sizes for each participant were divided into 10 deciles and were plotted along the X-axes. Panel *a* shows response time (z-scored) on the Y-axis. There was a U-shaped relationship for both distractor-present and distractor-absent trials. Panel *b*'s Y-axis shows the percentage of trials with the initial landing position on the distractor. Panel *c* shows the z-scores of the time needed to locate the target. Similar to response times, this measure also followed a U-shaped relationship with pupil size. Panel *d* shows the z-scored total looking time on the target. The measure did not follow a U-shaped relationship with pupil size. Therefore, both small and large pupil sizes were associated with

reduced task performance. Moreover, the increased response time at extreme levels of pupil size cannot be attributed to increased looking time on the target. Instead, it might be related to increased time searching for the target.

2.4.6 Pupil size and subjective reports of MW.

Our final analysis examined the relationship between pre-trial pupil size and subjective reports of MW. We did not find a significant difference between pupil size during on-task trials (Mean = $-.05$, SD = $.29$) and pupil size during MW trials (Mean = $.002$, SD = $.21$), $F(1, 63) = .83$, $p = .36$. The data are shown in Figure 2-7.

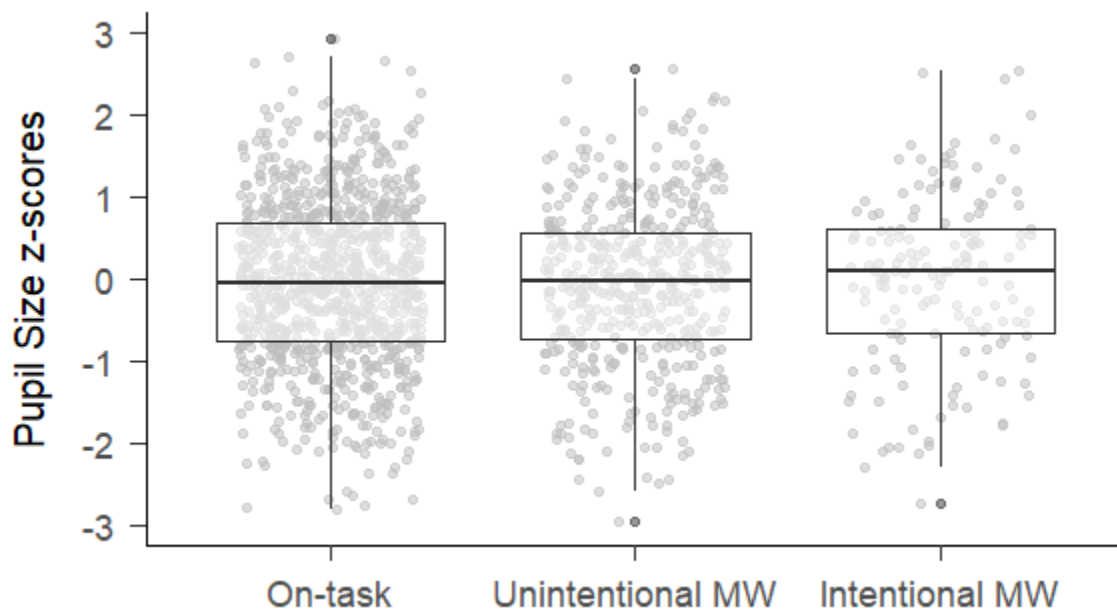


Figure 2-7. A boxplot showing pre-trial pupil size z-scores during trials classified as being on-task, unintentional MW, and intentional MW. Raw data points are plotted behind the boxplot.

2.5 Discussion

During visual search, it is important to avoid items that are irrelevant to the target. However, our focus on the task waxes and wanes, which may lead to fluctuations in task

performance. The current study examined if changes in attentional states would lead to deficits in visual search, and particularly increased distractor interference. Participants completed a visual search task to search for a specified target. A salient distractor occurred in a randomly selected 50% of trials and at a random location. At the behavioral level, we did find that participants took a longer duration to respond on distractor-present trials if they were mind-wandering, compared to when they were on-task. A similar pattern was not found for distractor-absent trials. Therefore, the presence of the distractor did somehow slow the participants down when they were mind-wandering. We then used eye-movement measures to pinpoint the locus of interference. We found that the interference cannot be attributed to an oculomotor capture by the salient distractor: participants could effectively suppress their first eye movements to the salient distractor across all trials (replicating Gaspelin et al. (2017)), and this oculomotor suppression effect did not vary as a function of self-reported attentional states. Participants also appeared to be equally fast at locating the target during mind-wandering and on-task episodes. Interestingly, participants spent longer time fixating the target if they were mind-wandering (compared to being on-task) on distractor-present trials but not on distractor-absent trials. Therefore, there was evidence of increased distractor interference during MW, although this interference was not apparent until the target had been located.

It is interesting to note that participants were able to suppress initial eye movements to the distractor even during MW. We reasoned that selection history might have played an important role in guiding attention in our task (Awh, Belopolsky, & Theeuwes, 2012). As typical in “feature-search mode” paradigms, participants were asked to search for a fixed target for many consecutive trials. When a stimulus is repeatedly selected on previous trials, it is more efficiently identified on the current trial; conversely, when a stimulus is repeatedly ignored, it tends to be

ignored on the current trial. Contrary to voluntary selection, history-based selection usually operates automatically and effortlessly (Theeuwes, 2018). Indeed, a recent study by Gaspelin, Gaspar, and Luck (2019) showed that the bias of selection history was so strong that it cannot be overridden by a voluntary selection of the target in the current trial. Therefore, participants might be still have been able to access their selection history of stimulus features during MW, and this allowed them to effectively suppress the salient distractor. On the other hand, this might be the very reason participants engaged in MW – they were so well-practiced to suppress the distractor that they believed engaging in MW would not hurt their performance. This notion is supported by recent evidence showing that participants could to some extent strategically regulate off-task thoughts without affecting primary task performance (Seli, Carriere, et al., 2018).

Another interesting finding is that total looking time on the target was longer on MW trials compared to on-task trials if the distractor was present. This effect did not seem to be caused by difficulties in retrieving the correct key response, because a similar pattern was not found for distractor-absent trials. One tentative explanation is that participants during MW needed a longer time to confirm that the correct item was selected if a distractor was present. Selection history may bias initial attention away from the distractor (and towards the target) in an automatic fashion, but during MW participants might need to “double-check” that this was indeed the item to which that they were supposed to respond (instead of the salient distractor). This prolonged checking was not necessary when there was no distractor present. This explanation is consistent with the observation that MW is often associated with goal neglect (Kane & McVay, 2012; McVay & Kane, 2009), although this was not reflected by a saliency-driven attentional capture.

The use of pupil size is motivated by the fact that pupil size covaries with activities in the LC, which plays an important role in balancing the trade-off between on and off-task focus. Consistent with the adaptive-gain theory of LC functions (Aston-Jones & Cohen, 2005), we found an inverted-U relationship between task performance (response time, in particular) and pre-trial pupil size, such that faster response times were achieved with intermediate levels of pupil sizes. Relatively small pupil sizes are usually associated with low-alertness and drowsiness, which is a likely reason for elevated pupil sizes. On the other hand, the increase of pupil size from intermediate to large usually marks the transition from exploitation (maintain focus on the current task) to exploration (disengage from the current task). In this state, participants' attention might be distracted for some reason so that they slowed down their response. It is worth noting that this inverted-U relationship was found for both distractor-present and distractor-absent trials, and there was little evidence of increased distractor interference at extreme levels of pupil size. In particular, we did not find that looking time on the target increased at extreme levels of pupil size as it did during self-reported MW. Therefore, while self-reported MW and extreme levels of pupil size were associated with increased response time, the underlying mechanisms might be different.

Contrary to some previous studies, we did not find a straightforward relationship between pre-trial pupil size and MW. From Figure 2-7, it was also apparent that pupil size at different attentional states followed similar distributions. Therefore, it seems unlikely that the lack of difference was due to MW being more likely to occur with both extremely large and extremely small pupil sizes. Unsworth and Robison (2018) proposed that the task nature can modulate the relationship between MW and pre-trial pupil size. In their experiment 1, they found that MW was associated with reduced pre-trial pupil diameter in a vigilance task; However, this

relationship disappeared when the participants were asked to generate a speech for an interview while doing the task (experiment 2). In their experiments 3 & 4, they further showed that the pupil size-MW relationship appeared when the stimulus onset asynchrony (SOA) was random or very short (2 secs), but not when the SOA was long (5 secs or 8 secs). Based on these findings, the authors suggested that the pupil size-MW relationship only appears when the task promotes external focus. In the context of our task, participants might not need much external focus to complete the task because they were well-practiced to locate the target. This might have reduced any relationship between pre-trial pupil size and self-reported MW.

It is important to note that we only tested the pre-trial (tonic) pupil size in our study. Previous studies examining the relationship between task-evoked pupillary response (phasic pupil size) found that MW was consistently associated with smaller task-evoked responses (e.g., Jubera-García et al., 2019; Mittner et al., 2014; Unsworth & Robison, 2016, 2018). An important difference between our task and theirs is that our task involves making overt and fast eye movements away from the screen center, which may occlude any effects on pupil dilation. However, we do note that there are ways to correct this artifact (Hayes & Petrov, 2016). This remains an interesting line of research for future studies.

The relationship between eye movements and MW has been primarily studied in reading comprehension (Faber et al., 2018; Foulsham et al., 2013; Reichle et al., 2010; Steindorf & Rummel, 2019). The current study extended this line of research by showing that self-reported MW was also associated with performance deficits in a visual search task. We hope that the eye-tracking results would be helpful to the efforts of building automatic detectors of MW that is based on eye movement measures. Moreover, we showed that pre-trial pupil size seemed to be a valid measure of task engagement as it was associated with changes in certain performance

metrics. However, we did not find a straightforward relationship between subjective and objective measures of attention. This finding may complicate the hope of using pre-trial pupil size as an objective indication of MW. Rather, changes in pupil size and self-reported attentional states might reflect different forms of task disengagement, and future research is needed to disentangle their relationships.

Chapter 3 Mind-wandering during Real-world Scene Perception

3.1 Abstract

Recent studies have shown that mind-wandering (MW) is associated with changes in eye movement parameters. The current study examined eye movement features of MW during scene perception, focusing on two aspects of eye movements that received little attention in the previous literature. The first is the sequential pattern of eye movements, namely scanpaths, during MW. The second is how eye movements relate to scene properties during MW. The current study aimed to address both by (1) conducting a Recurrence Quantification Analysis (RQA; Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013) of eye movements during MW, and (2) examining how fixations during MW are related to scene meaningfulness (Henderson & Hayes, 2017). Participants completed a real-world scene encoding task and responded to thought probes assessing intentional and unintentional MW. Both types of MW were associated with worse memory of the scenes. Importantly, RQA showed that scanpaths during unintentional MW were more repetitive than when on-task, as indicated by a higher recurrence rate and more stereotypical fixation sequences. This increased repetitiveness suggests that the viewer had to refresh previous locations more frequently to remember the information. We also found that fixation allocation during intentional MW was less related to the distribution of meaning across the scene. Finally, we were also able to validate several traditional measures: both intentional and unintentional MW were associated with fewer and longer fixations; Eye-blinking increased numerically during both types of MW but the difference was only significant for unintentional

MW. Overall, the results advance our understanding of how visual processing is affected during MW by highlighting the sequential and local characteristics of eye movements.

3.2 Introduction

Every so often, we get absorbed by our thoughts that are not related to the here and now. Mind-wandering (MW) refers to self-generated thoughts that are irrelevant to the current task. MW is often thought to involve “perceptual decoupling”, or disengagement of attention from external information and events (Schooler et al., 2011). Previous work has shown that MW is associated with reduced cortical processing of visual input (Braboszcz & Delorme, 2011; Kam et al., 2011; also see Mo, Liu, Huang, & Ding, 2013). One typical way of obtaining MW data is through the use of thought probes. In this paradigm, a thought probe is occasionally presented during an external task to inquire about participants’ current thoughts and focus. Using the probe-caught method, studies have shown that MW is associated with reduced performance in tasks that require sustained attention (for a review, see Mooneyham & Schooler, 2013). Studies have also shown that, when the mind wanders, eye movements change as well. For example, while mind-wandering during reading, readers produced fewer fixations (Uzzaman & Joordens, 2011), read more slowly (Foulsham et al., 2013), and their fixations were less associated with information presented in the text (Reichle et al., 2010). MW was also found to be associated with increased eye-blinking (Smilek et al., 2010). In particular, one previous study examined eye movement signatures of MW in a scene perception task (Krasich et al., 2018). They found that probe-caught MW was associated with fewer and longer fixations, as well as greater dispersion (fixations located further away from each other). The link between eye movements and MW has also inspired attempts to develop gaze-based detectors of MW (e.g., Bixler & D’Mello, 2016; Faber et al., 2018).

Despite these fruitful findings, two aspects of eye movements have received little attention in the current literature of MW. The first one is the sequential aspect of eye movements, namely scanpaths. The second one is the role of top-down and bottom-up factors in guiding eye movements during MW. Below, I discuss how studying each aspect can advance our understanding of visual processing during MW.

3.2.1 Scanpath analysis.

Fixations naturally unfold over time and simply aggregating them may lose important information about online cognitive processing. Early work by Yarbus (1967) showed that scanpaths on the same painting unfolded differently depending on the task goal. Noton and Stark (1971b, 1971a)'s *scanpath theory* further proposed that scanpaths are an external reflection of the viewer's internal cognitive representation. Since then, a growing literature has employed scanpath analysis to study online cognitive processing during scene perception (Foulsham & Underwood, 2008), visual search (Dewhurst et al., 2018), reading (von der Malsburg & Vasishth, 2013), social communication (Richardson & Dale, 2005), domain expertise (McIntyre & Foulsham, 2018), etc. Meanwhile, the field has seen a large growth in the number of new methods to analyze different aspects of scanpaths (for a review, see Anderson, Anderson, et al., 2015).

Analyzing scanpaths may uncover novel information about how attentional selection and eye movements are affected during MW. There is extensive research looking at how bottom-up and top-down factors affect attention during scene perception (e.g., Henderson, 2003; Tatler, Hayhoe, Land, & Ballard, 2011). But most of the studies assumed that participants were actively trying to process the scene with their complete focus. This is an optimistic assumption in either

experimental or daily life settings, as previous studies have shown (Kane et al., 2007; Seli, Risko, Smilek, et al., 2016). MW is an interesting case because it is a top-down effect that is completely unrelated to the scene itself. How do eye movements unfold differently when the viewer is thinking about something else? Answering this question can help us understand the control processes that shape visual processing under different attentional states. Given the prevalence of MW, it may also have practical implications because scene perception is a fundamental part of daily activities.

The current work will focus on one aspect of scanpaths: the pattern of refixations (i.e., fixations returning to previously viewed locations). Refixations are an important part of normal gaze behavior across different tasks and have been shown to play a functional role in remembering and comprehending information (e.g., Meghanathan, Nikolaev, & van Leeuwen, 2019; Schotter et al., 2014; Zelinsky, Loschky, & Dickinson, 2011). Recently, Anderson et al. (2013) introduced Recurrence Quantification Analysis (RQA) to capture the temporal pattern of refixations. RQA is a method to understand the temporal structure of a single scanpath (unlike other scanpath measures that require comparisons between two different sequences of eye movements). This feature makes RQA measures relatively easy to interpret and is convenient for statistical testing. The calculation of RQA measures is also independent of the specific content of the stimuli, making them potential candidates of task-general predictors of MW (cf. Faber et al., 2018).

Two fixations are considered recurrent if their physical distance is below a pre-defined threshold (usually 1° - 2° of visual angle corresponding to foveal vision; Anderson et al., 2013). For a given fixation sequence, we can illustrate all recurrent fixations at all time lags on a plot like the one in the panel *a* of Figure 3-1. In this recurrence plot, the X- and Y-axis represent the

same fixation sequence, with the numbers representing the fixation index (e.g., 1st fixation, 9th fixation, etc.). Recurrent fixations are denoted as a black dot. Each fixation is recurrent with itself, as shown by the major diagonal line. Above the line of self-recurrence, we can see multiple cases where later fixations are recurrent with earlier fixations (e.g., fixations 11, 19, and 20 are recurrent with fixation 1).

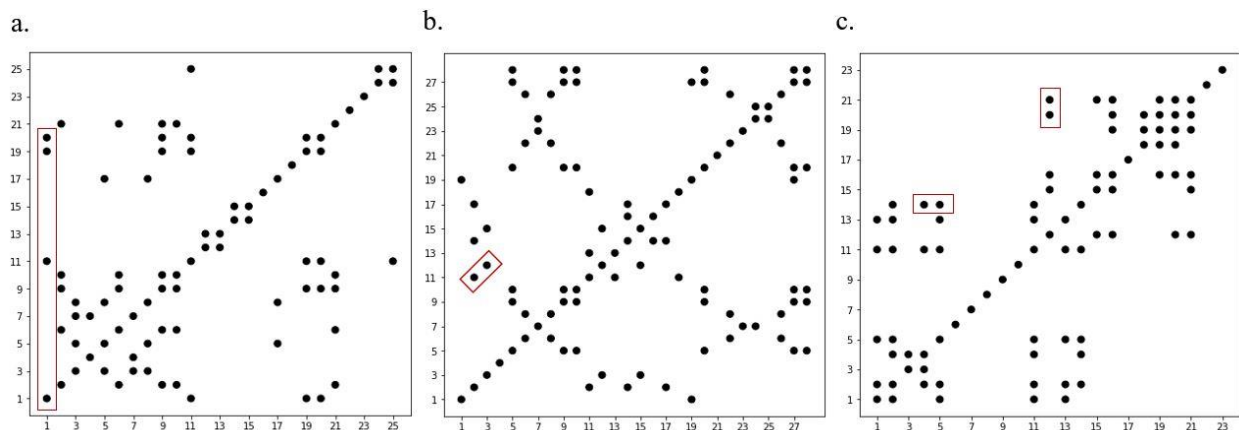


Figure 3-1. Example of recurrence plots. The X- and Y-axis represent the same fixation sequence, with the numbers representing the fixation index (e.g., 1st fixation, 9th fixation, etc.). Two fixations are considered recurrent (denoted as a black dot) if their physical distance is below a pre-defined threshold (usually 1° - 2° of visual angle corresponding to the foveal vision; Anderson et al., 2013). The red rectangle in Panel *a* illustrates that fixations 11, 19, and 20 are recurrent with fixation 1. Panel *b* illustrates determinism. As highlighted in the red rectangle, areas examined by fixations 2 and 3 are re-examined later in this trial in the same order by fixations 11 and 12. Panel *c* illustrates laminarity. As highlighted in the red rectangles, an area examined by fixation 12 is later re-examined by fixations 20 and 21, forming a vertical line. In contrast, an area initially examined by fixations 4 and 5 are later re-examined by fixation 14, forming a horizontal line. Finally, compared to panel *b*, the recurrent fixations in panel *c* are closer to the major diagonal, meaning that refixations in panel *c* occur closer together in the fixation sequence. Therefore, the center of recurrence mass (CORM) is smaller in panel *c* than in panel *b*.

From the recurrence plot, Anderson et al. (2013) identified and described four recurrence measures which are briefly outlined here (also see Gurtner, Bischof, & Mast, 2019). First, the *recurrence rate* indicates the percentage of recurrent fixations for a given fixation sequence. It indicates how often a viewer returns to previously inspected areas. Second, *determinism* indicates the percentage of recurrent points that fall on diagonals parallel to the line of self-recurrence (as illustrated in Figure 3-1, panel *b*). Determinism increases when a viewer moves the gaze from area A to B to C and later repeats the same sequence in the same order. Therefore,

determinism represents repeated fixation sequences. Third, *laminarity* indicates the percentage of recurrent fixations that fall on horizontal or vertical lines in the recurrence plot (as illustrated in Figure 3-1, panel *c*). Vertical lines represent cases where an area is first examined by a single fixation and later on examined by consecutive fixations. In contrast, horizontal lines represent cases where an area is first examined and later on re-examined with a single refixation. For determinism and laminarity, usually, a minimum line-length threshold of 2 is set to rule out single refixations. Fourth, *the center of recurrence mass* (CORM) measures the overall position of recurrence points in relation to the line of self-recurrence. A small CORM value indicates that refixations occur relatively close in time (i.e., close to the line of self-recurrence). For example, if a viewer tends to immediately re-examine a previous location, the recurrent points will be close to the line of self-recurrence, producing a small CORM value. Alternatively, if a viewer tends to re-examine a previous location after many intermediate fixations at other locations, the recurrence points will be far from the line of self-recurrence, producing a large CORM value.

To summarize, the recurrence rate (overall percentage of refixations) and CORM (whether refixations occur soon or late) capture the global pattern of recurrences in a scanpath. Determinism (the percentage of repeated fixation sequences) and laminarity (the percentage of detailed inspections) further characterize the local pattern of recurrences.

How would the pattern of refixations change during MW? One possibility is that the viewer tends to skim over the entire picture without detailed processing at previous locations. This viewing pattern may lead to a reduced overall recurrence rate, and changes in the local structures (e.g., reduced determinism and laminarity) as well. Looking at new rather than old regions might have the benefit of gathering as much information as one possibly could, given the limited attentional resources allocated to the external task during MW. Krasich et al. (2018)

reported that fixations during MW were located farther away from each other, although this effect was only significant for one 5-second bin. This finding seems to suggest that fixations would be less recurrent during MW. Alternatively, one can predict that scanpaths are *more* repetitive during MW. Deficits in visual processing may require the viewer to perform more refixations than usual to effectively process visual information. Refixations are known to be caused by insufficient processing of information at previous locations (Droll & Hayhoe, 2007; Gilchrist & Harvey, 2000; Meghanathan et al., 2019; Zelinsky et al., 2011). Gilchrist and Harvey (2000) noted that refixations “constitute a failure of memory”. Zelinsky et al. (2011) further propose that refixations serve a rehearsal function to refresh fading object representations in visual working memory. Thus, fixations might be more recurrent during MW as a way to compensate for processing deficits. These two possibilities are opposite to one another, but both seem to be sensible eye movement patterns in response to the reduced visual processing during MW. Notably, aggregated eye movement measures do not permit distinguishing between these two predictions, but scanpath analysis can reveal how MW affects the way that looking unfolds over time.

3.2.2 The role of scene meaningfulness and saliency.

Although RQA can capture the intricacies of a scanpath, it does not take into account the specific context in the scene. Eye movements usually reflect specific features of the task (e.g., novelty, difficulty, etc.) but this association may cease to exist during MW given the reduced external focus. For example, Reichle et al. (2010) found that the robust association between word frequency and eye movements usually observed during normal reading disappeared during mindless reading. These results have been taken as evidence of reduced attentional coupling with the external task during MW.

As with reading, eye movements during scene perception are often associated with the properties of the scene. Much of the MW research has been built on the idea that MW involves a state of “attentional decoupling”, or superficial processing of external information (Schooler et al., 2011). The attentional decoupling account of MW would predict that eye movements during scene perception are less associated with specific content in the scene. While this seems to be a reasonable and even obvious hypothesis, to our knowledge no published study so far has explicitly examined this. One previous study (Krasich et al., 2018) did examine eye movement patterns of MW during scene perception, but the measures examined were not related to specific content in the scene. Therefore, it is critical to analyze if the relationship between eye movements and the scene would change during MW.

There is an ongoing debate about whether visual attention is driven by bottom-up or top-down factors (e.g., Gaspelin & Luck, 2018b; Henderson, 2003; Henderson, Malcolm, & Schandl, 2009; Itti & Koch, 2000; Theeuwes, 2010; Yantis & Jonides, 1990). Some have argued that visual attention operates in a primarily bottom-up manner such that attention is automatically captured by visually salient regions of the scene (e.g., Itti & Koch, 2000; Parkhurst, Law, & Niebur, 2002; Theeuwes, 2010). Saliency maps, as a computational implementation of this idea, quantify the distribution of saliency across the entire scene (e.g., Harel, Koch, & Perona, 2007; Itti & Koch, 2000). Regions that are more distinctive or salient relative to others receive higher saliency values. However, empirical studies have shown that saliency maps are only modestly successful at predicting actual fixations in real-world scenes (Tatler et al., 2011). On the other hand, there is a large body of research showing that top-down influences on visual attention can completely override predictions based on saliency (e.g., Anderson, Ort, et al., 2015; Bacon & Egeth, 1994; Gaspelin et al., 2017; Henderson, Brockmole, Castelhana, & Mack, 2007;

Henderson & Hayes, 2017; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006). Recently, Henderson and Hayes (2017) showed that scene meaningfulness is better at predicting fixation location than saliency. Specifically, they cut each scene into a large number of small patches and asked people how each patch is informative or recognizable. These individual ratings were averaged, smoothed, and combined to produce a “meaning map” for each scene. Analogous to saliency maps, a meaning map represents which regions in the scene are relatively more meaningful. Finally, they obtained the ground-truth attention maps by asking an independent sample to study these scenes. They found that the attention map was more similar to the meaning map than the saliency map produced by a prominent saliency model (Graph-Based Visual Saliency (GBVS); Harel et al., 2007). This finding was replicated even when the viewer was not required to process the meaning of scenes (e.g., rating the brightness of scenes; Peacock, Hayes, & Henderson, 2019). These results, according to the authors, indicate that attention during real-world scene viewing is primarily driven by meaning rather than saliency.

As an extension of Henderson and Hayes (2017)’s finding, we wanted to test if the degree of which attention is associated with meaning in real-world scene perception depends on the viewer’s attentional state. We hypothesize that the association with meaning should be stronger for attention maps based on “on-task” fixations (i.e., those obtained from on-task episodes), compared to those based on MW fixations.

If attention is less associated with meaning during MW, would it be more associated with saliency instead? Deficits in the cognitive processing of meaning may cause eye movement control to degenerate such that it operates in a purely bottom-up manner. In other words, during MW, eye movements might be driven by whatever is salient albeit not meaningful in the scene. Because meaning and saliency can be highly correlated for natural scenes (Henderson, 2003), the

answer to this question lies in the *unique* variance that can be accounted for by meaning and saliency (i.e., holding the confounding variable constant statistically). Henderson and Hayes (2017) showed that, compared to saliency maps, meaning maps also accounted for more *unique* variance of the actual eye movements in their study. However, if MW involves qualitatively changes in the role of meaning and salience in guiding attention, the advantage of meaning over saliency as observed in their study may disappear or even reverse.

3.2.2 The current study.

In the current study, we asked participants to study a range of real-world scenes for a later memory test. Thought probes were occasionally presented to inquire about participants' task focus. MW is an umbrella term that encompasses a wide range of mental phenomena (Seli, Kane, Smallwood, et al., 2018). For example, recent studies have shown that MW can occur with or without intention (Seli, Risko, Smilek, et al., 2016). Unintentional MW occurs when individuals lose control over their thoughts despite their best intention to focus on the task, whereas intentional MW occurs when individuals decide to disengage from the task. Previous research showed that these two types of MW are associated with different psychological and neurological processes (Phillips et al., 2016; Seli, Risko, & Smilek, 2016; Seli, Smallwood, et al., 2015). Therefore, we measured both intentional and unintentional MW.

Data analysis consisted of three parts:

First, we examined the four previously described RQA measures and how these measures might change when a viewer is on-task compared to when they are intentionally or unintentionally mind-wandering.

Second, we re-examined several conventional eye movement measures previously shown to be associated with MW, including fixation count, duration, and eye-blinks. Specifically, previous studies showed that MW was associated with fewer but longer fixations (e.g., Faber et al., 2018; Foulsham et al., 2013; Krasich et al., 2018). Smilek et al. (2010) reported that readers blinked more often when they were mind-wandering than when they were on-task during reading comprehension. The authors interpreted the increased blinking as an overt embodiment of perceptual decoupling. However, subsequent studies obtained mixed results (Faber et al., 2018; Grandchamp et al., 2014; Krasich et al., 2018; Uzzaman & Joordens, 2011). Therefore, it is worth examining this measure again with improved detection methods and statistical power. It also remains an open question whether intentional and unintentional MW are associated with similar eye movement patterns.

Third, we examined the role of meaning and saliency in guiding eye movements during intentional and unintentional MW. We constructed meaning maps as described in (Henderson & Hayes, 2017) to represent how meaning was distributed across each scene. Based on the attentional decoupling account of MW (Schooler et al., 2011), we hypothesized that attention should be less associated with meaning during MW; we also wanted to test if saliency maps (as defined in Harel et al., 2007), compared to meaning maps, play a larger role in guiding eye movements during MW.

3.3 Method

3.3.1 Participants.

The statistical power of the current design is related to the number of participants, and the number of probes, as well as the actual MW rate. The last of these factors are not under the

control of the experimenters. For multilevel modeling with an unbalanced design, there is currently no consensus on power analysis. Previous studies who examined similar measures also vary substantially in their design (Foulsham et al., 2013; Krasich et al., 2018; Reichle et al., 2010; Smilek et al., 2010; Steindorf & Rummel, 2019; Uzzaman & Joordens, 2011). Given this situation, we decided to collect a larger number of total trials (participants*probes) than previous studies have typically done. In particular, one previous study on scene perception recruited 51 participants * 8 thought probes (Krasich et al., 2018, exp. 1); They were able to replicate their main findings with a smaller number of trials in Experiment 2. The increased blinking during MW was first reported in a study with 15 participants * 10 probes (Smilek et al., 2010). Steindorf and Rummel (2019) examined MW during a reading task with 122 participants * 10 probes, which is one of the largest-sample studies so far. Our study recruited 64 undergraduate students (Age: *Mean* = 18.84, *SD* = .79; 64% female; all with normal eyesight) from the University of Michigan to participate for course credit. We decided to present 36 thought probes for each participant. After discarding data from 7 participants who had low tracking ratios (< 75%, a threshold set *a priori*), the final sample size is 57 participants.

3.3.2 Apparatus and stimuli.

3.3.2.1 Pictures.

We selected 180 pictures as the to-be-studied material from the SUN database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010) and the LabelMe database (Russell, Torralba, Murphy, & Freeman, 2008). The set consisted of 60 exteriors, 60 interiors, and 60 landscapes. We selected only pictures without identifiable human faces. Of these pictures, we randomly selected 12 of each type as the target pictures. These target pictures were the same across participants,

and each target picture was followed by the presentation of a thought probe (discussed below). The order of presentation was randomized with the constraint that any two target pictures were separated by at least three non-target pictures. Another 72 pictures from the same database (24 for each type) were selected as new pictures in the memory test. All pictures were presented in 1024*768 pixels on a 20.1-inch screen at approximately 70 cm to the participant. Thus, each picture subtended about 32°*25° of visual angle.

3.3.2.2 Thought probes.

A thought probe was presented after every target picture, asking “*where was your attention during the last picture?*”. Participants chose either “*I was focusing on the picture*” by pressing “A” on the keyboard or “*I was thinking about something else*” by pressing “B” on the keyboard. If the latter was chosen, participants were further asked to indicate their intentionality of MW by pressing “A” for intentional MW and “B” for unintentional MW. “On-task” was communicated to the participants as “*you were focused on completing the task and were not thinking about anything unrelated to the task*”; MW was communicated to the participants as “*you were thinking about something completely unrelated to the task*”, with intentional MW as “*you intentionally decided to think about things that are unrelated to the task*” and unintentional MW as “*your thoughts drifted away despite your best intentions to focus on the task*” (Seli, Cheyne, et al., 2015).

3.3.2.3 Eye-tracking.

Monocular eye movements were recorded by an Eyelink 1000 tracker at a sampling rate of 500 Hz. No chin rest was used, and head movement was adjusted by tracking a sticker on the participant’s forehead. The experiment was implemented by the OpenSesame software (Mathôt,

Schreij, & Theeuwes, 2012) with functions from the PyGaze package (Dalmaijer, Mathôt, & der Stigchel, 2014).

3.3.3 Procedure.

After signing the informed consent form, participants were asked to study pictures for a memory test. Participants learned that a thought probe would occasionally occur after some of the pictures and ask them whether they were mind-wandering or on-task during their viewing of the picture just presented. Participants first completed a practice block that consisted of 5 example trials (with a thought probe) and a recall test. Then, the experimental blocks began. Exteriors, interiors, and landscapes were presented in three separate blocks. The block order was counterbalanced across participants. Each block had a study-test structure. In the study phase of each block, participants studied 60 pictures consecutively. Each picture was presented for 10 seconds. The structure of each trial was shown in Figure 3-2. A thought probe was presented after every target picture. In the test phase of each block, participants indicated whether each picture was an old one or a new one from a set of 12 target pictures, 12 randomly selected non-target pictures, and 24 new pictures of the same type. The eye-tracker was calibrated using a 5-point calibration before each block. Participants were asked to focus on a white dot appearing at the top, bottom, left, right, and the center of the screen while the experimenter manually confirmed the fixated locations by pressing the spacebar on the host computer. Then, a validation procedure was performed to evaluate the spatial errors of the calibration. Eye movements were recorded during both the study and the test phase but only data from the study phase were analyzed and reported.

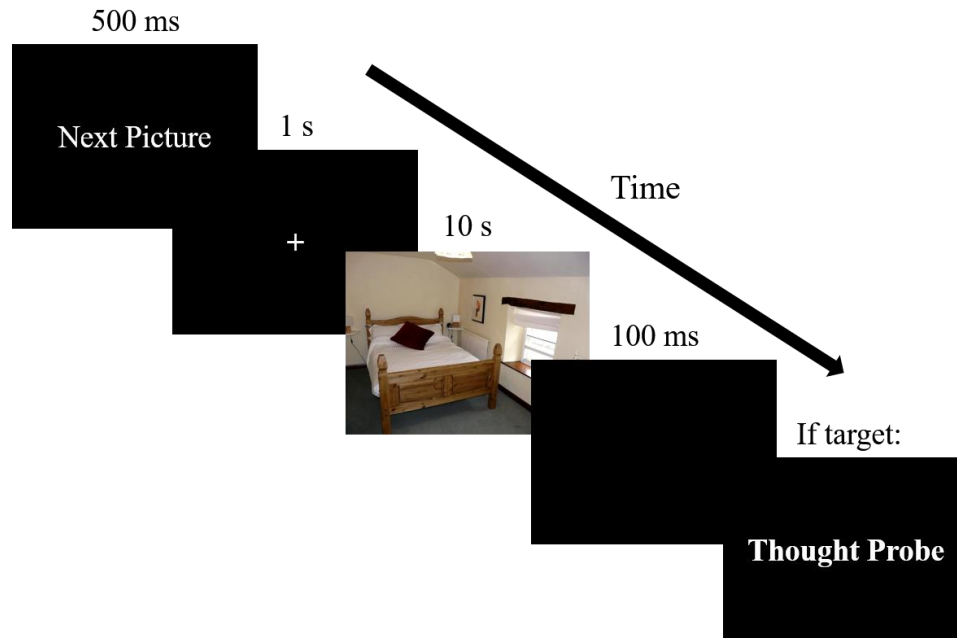


Figure 3-2. An example trail during the study phase. A trial started with a 500-millisecond display of “Next Picture”, followed by a 1-second fixation cross. Then the picture appeared for 10 seconds, followed by a black screen for 100 milliseconds. For target trials, a thought probe would then appear to ask whether participants were mind-wandering during the picture they just saw.

3.3.4 Calculation of RQA measures.

Data analysis was conducted in R (Version 3.6.1; R Core Team, 2018). We discarded fixations greater than 2000 msec or shorter than 80 msec (8.1% of data) and fixations outside of the screen (1.82% of data) before computing eye movement measures. Trials with no fixations (1.82% of trials) were also discarded. Recurrence measures were computed based on a radius of 60 pixels (about 2° of visual angle; Anderson et al., 2013). A threshold of 2 consecutive recurrences (as default) was set to rule out single refixations in the calculation of determinism and laminarity. The code for computing RQA measures can be found at <https://barlab.psych.ubc.ca/research/>. Note that for determinism, laminarity, and CORM, zeroes were treated as missing per the program’s default. See the supplemental material for the number of trials analyzed for each measure.

RQA measures were analyzed by linear mixed effect models (Bates et al., 2015). Attention (intentional MW/unintentional MW/on-task) was dummy-coded with “on-task” being the reference level. Picture type (exteriors/interiors/landscapes) entered the model as a covariate. We started by fitting all models with all possible random intercepts and slopes (Barr et al., 2013). In the case of non-convergence or singular fit, we performed backward elimination of the random effects using the *step()* function from the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017). To obtain *p*-values, we used the Satterthwaite approximation and further adopted the Holm-Bonferroni correction. In the main text, we report the difference between MW and on-task episodes, aggregated across picture types. The full model results can be accessed at <https://osf.io/6pj9m/files/>.

3.3.5 Map creation.

3.3.5.1 Attention maps.

For each picture, we pooled fixations across participants and grouped them based on thought probe responses. Attention maps were created by first building a two-dimensional fixation frequency matrix and then smoothing the matrix using a two-dimensional Gaussian filter (window size: 120*120 pixels, standard deviation: 30 pixels). This procedure would produce three attention maps (on-task, intentional MW, and unintentional MW) for each picture. Note that no intentional MW was reported for two of the target pictures. Thus, for these two pictures, no intentional MW map was generated. The code is available at <https://github.com/HanZhang-psych/SceneMeaningMapping>. Attention map examples are shown in Figure 3-3, panels *b-d*.

3.3.5.2 *Meaning maps.*

Meaning maps were created following the same procedure as described in Henderson and Hayes (2017). Each of the 36 target pictures was cut into a series of partially overlapping circular patches. For each picture, three-hundred fine patches (diameter = 87 pixels) and 108 coarse patches (diameter = 205 pixels) were extracted, producing 10800 unique fine patches and 3888 unique coarse patches (14688 patches in total). 147 workers from Prolific.co, an online crowdsourcing platform, rated these patches for \$3 of participation reward. We set the following criteria for participation: (1) Age: 18 - 35, (2) Nationality: US or Canada, (3) Fluent language: English, (4) Approval rate: $\geq 99\%$, (5) Number of previous submissions: ≥ 100 . Workers were asked to rate how “meaningful” they thought each scene patch was using a 6-point scale (1 - very low, 6 - very high). Each worker rated 300 unique patches, so each patch was rated 3 times by 3 unique workers. Patches were presented without the scene context.

Rating scores were mapped back to each pixel in the original pictures. For example, if a patch was rated as “2”, then all pixels covered by this patch would include a score of 2. Note that due to the overlapping nature of the patches, some pixels (e.g., those in the middle of a picture) included more ratings than others (e.g., those at the edges of a picture). The meaning map was created by averaging and smoothing the rating scores at each pixel. As in Henderson and Hayes (2017), we applied a central bias (i.e., image center tend to receive more attention than image peripherals) to the meaning maps to improve predicting power. This central bias (derived from the “invCenterBias.mat” in GBVS toolbox) was applied to the raw meaning map using pixel-wise multiplication. The code for generating the meaning maps can be found at the original authors’ website: <https://osf.io/654uh/>. An example of the meaning maps is shown in Figure 3-3, panel *e*.

3.3.5.3 Saliency maps.

Saliency maps for the target pictures were generated using the GBVS toolbox (Harel et al., 2007) with default parameter settings. GBVS is a prominent saliency model and was used in Henderson and Hayes (2017) to evaluate the meaning maps. Central bias is an inherent feature of GBVS so the bias matrix applied to meaning maps was not applied here. The toolbox can be found at the original authors' website at <http://www.vision.caltech.edu/~harel/share/gbvs.php>. An example of the meaning maps is shown in Figure 3-3, panel *f*.

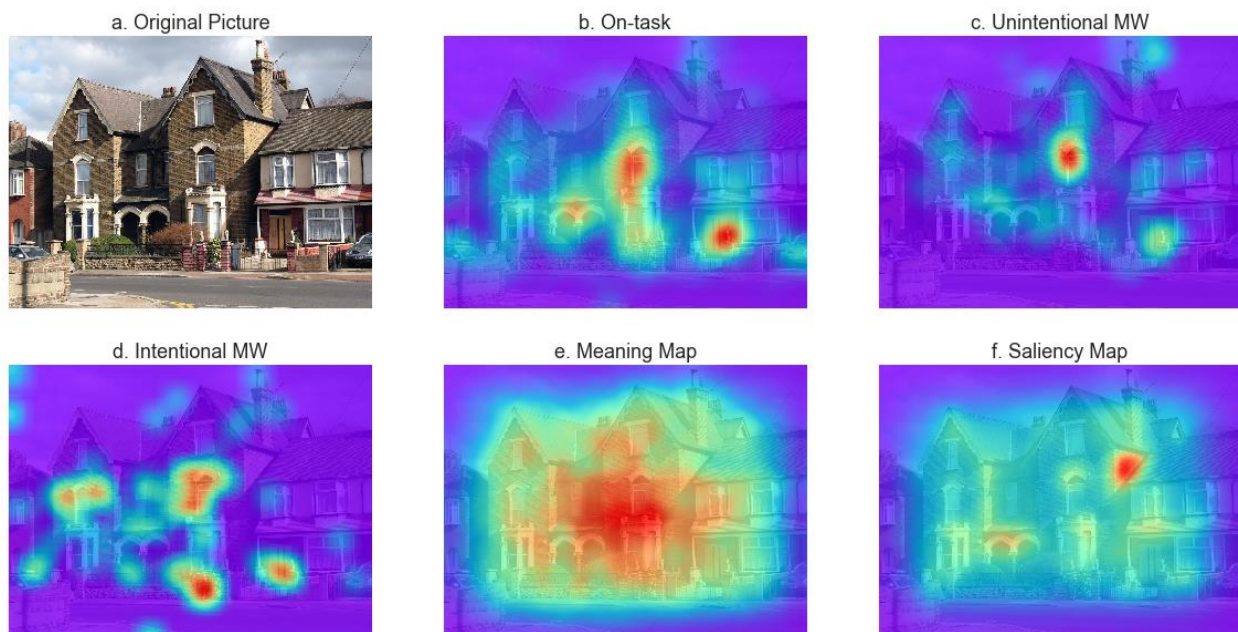


Figure 3-3. Maps created for one of the target pictures. Panel *a* shows the original picture. Panel *b* shows the attention map generated from “on-task” fixations across participants. Panel *c* shows the attention map generated from “unintentional MW” fixations across participants. Panel *d* shows the attention map generated from “intentional MW” fixations across participants. Panel *e* shows the meaning map, following procedures in Henderson and Hayes (2017). Panel *f* shows the saliency map generated by the GBVS toolbox with default parameters.

3.4 Results

It is worth noting that responses to the thought probes were heavily unbalanced: participants on average reported being on-task for about 73% of the time ($SD = 22\%$), followed by unintentional MW ($Mean = 22\%$, $SD = 19\%$), and intentional MW ($Mean = 5\%$, $SD = 8\%$).

The overall MW rate is consistent with previous literature ($Mean = 27\%$, $SD = 22\%$ in Krasich et al., 2018), but the rate of intentional MW was unexpectedly low (e.g., Seli, Carriere, et al., 2018; Seli, Cheyne, et al., 2015; Seli, Risko, & Smilek, 2016). Therefore, we suggest the reader be cautious when interpreting results regarding intentional MW due to limited observations.

3.4.1 Recognition of target pictures.

If participants were unintentionally mind-wandering rather than on-task during the presentation of target pictures, they were less likely to recognize those pictures during the memory test, $b = -0.15$, $SE = 0.04$, $t = -3.89$, $p < .001$. Similarly, participants had worse recognition if they were intentionally mind-wandering during the presentation of target pictures, $b = -0.15$, $SE = 0.06$, $t = -2.44$, $p = .020$.

3.4.2 Recurrence rate.

The recurrence rate indicates the proportion of fixations directed to previously inspected areas. We found that recurrence rate was significantly higher during unintentional MW than during on-task episodes, $b = 1.79$, $SE = 0.64$, $t = 2.80$, $p = .015$. During intentional MW, recurrence rate was also numerically higher compared to on-task episodes, but this difference was only marginally significant, $b = 3.50$, $SE = 2.03$, $t = 1.72$, $p = .097$. These results were demonstrated in Figure 3-5, panel *a*.

3.4.3 Determinism.

Determinism indicates repeated fixation sequences. Unintentional MW was associated with a significantly higher determinism compared to on-task episodes, $b = 4.03$, $SE = 1.13$, $t =$

3.55, $p = .001$. However, a similar pattern was not found for intentional MW, $b = -2.10$, $SE = 2.04$, $t = -1.03$, $p = .303$. Results of determinism were shown in Figure 3-5, panel *b*.

3.4.4 Laminarity.

Laminarity represents areas initially examined with a single fixation and later on followed-up with consecutive fixations or *vice versa*. There was no significant difference in laminarity between unintentional MW and on-task episodes, $b = 1.04$, $SE = 0.93$, $t = 1.12$, $p = .481$, or between intentional MW and on-task episodes, $b = 2.02$, $SE = 1.72$, $t = 1.17$, $p = .481$. These results were shown in Figure 3-5, panel *c*.

3.4.5 Center of recurrence mass (CORM).

The center of recurrence mass (CORM) indicates the whether refixations occur close in time. We did not find any significant difference between MW and on-task episodes in CORM. Unintentional MW: $b = 0.07$, $SE = 0.56$, $t = 0.12$, $p = .901$; intentional MW: $b = 1.37$, $SE = 1.02$, $t = 1.34$, $p = .358$. These results were shown in Figure 3-5, panel *d*.

To further illustrate the scanpath patterns, we plotted representative scanpaths during unintentional MW and on-task episodes in Figure 3-4. In this plot, each red point represents a single fixation. Recurrent fixations are connected by a black line. As shown in the figure, the scanpath during unintentional MW (on the right) was more repetitive compared to the on-task scanpath (on the left).



Figure 3-4. Representative scanpaths during unintentional MW and on-task episodes. Each red point represents a single fixation. Recurrent fixations are connected by a black line. As shown in the figure, the scanpath during unintentional MW (on the right) was more repetitive compared to the on-task scanpath (on the left). A high-quality version of this figure can be found at <https://osf.io/6pj9m/files/>.

3.4.6 Fixation dispersion.

The increased recurrence of scanpaths during MW, as shown by RQA, may lead to changes in the overall spatial distribution of fixations. We calculated fixation dispersion to indicate how fixations were spread across the stimulus (Krasich et al., 2018). It can be computed as the root mean square of the Euclidean distance from each fixation to the average position of all fixations. It is reported on a 0-1 scale by normalizing on the maximum dispersion possible, with higher values indicating greater dispersion.

Fixation dispersion and recurrence rate were significantly correlated, $r = -0.57$, $t = -31.17$, $df = 1981$, $p < .001$. Fixation dispersion was significantly smaller during unintentional MW than during on-task episodes, $b = -0.02$, $SE = 0.005$, $t = -3.86$, $p < .001$. For intentional MW, fixation dispersion was also smaller compared to on-task episodes, $b = -0.02$, $SE = 0.01$, $t = -2.95$, $p = .003$. Results are shown in Figure 3-5, panel *e*.

3.4.7 Area covered.

We computed another measure to indicate the spatial property of fixations. Specifically, we “drew” a circle with a radius of 60 pixels (about 2° of visual angle) around each fixation and calculated the total area covered by these regions in proportion to the full picture size. Therefore, this measure roughly indicated the percentage of area covered by foveal vision.

Recurrence rate was significantly correlated with the percentage of area covered, $r = -0.63$, $t = -36.55$, $df = 1981$, $p < .001$. Unintentional MW was associated with less coverage compared to on-task episodes, $b = -0.02$, $SE = .003$, $t = -6.96$, $p < .001$. A similar relationship was found for intentional MW, $b = -0.03$, $SE = 0.01$, $t = -4.86$, $p < .001$. Results are shown in Figure 3-5, panel *f*.

3.4.8 Fixation count.

Unintentional MW was associated with fewer fixations compared to on-task episodes, $b = -2.62$, $SE = 0.35$, $t = -7.51$, $p < .001$. A similar relationship emerged for intentional MW, $b = -3.39$, $SE = 0.63$, $t = -5.40$, $p < .001$. Results are shown in Figure 3-5, panel *g*.

3.4.9 Fixation duration.

The mean duration of fixations was longer during unintentional MW than during on-task episodes, $b = 24.81$, $SE = 9.76$, $t = 2.54$, $p = .028$. Similarly, the mean duration was longer during intentional MW than during on-task episodes, $b = 53.75$, $SE = 22.50$, $t = 2.39$, $p = .028$. Results are shown in Figure 3-5, panel *h*.

3.4.10 Eye-blinking.

We detected eye-blinking using a novel algorithm developed by Hershman, Henik, and Cohen (2018). The algorithm detects eye blinks based on the change in pupil size between subsequent samples and was shown to be more accurate than some other currently used methods. To remove artifacts, we defined a “normal” blink to be between 50 ms and 500 ms (Wang, Toor, Gautam, & Henson, 2011) and removed observations outside this range. We found that unintentional MW was associated with more blinks compared to on-task episodes, $b = 0.28$, $SE = 0.10$, $t = 2.79$, $p = .011$. Intentional MW had numerically more blinks than on-task episodes but the difference was not statistically significant, $b = 0.30$, $SE = 0.27$, $t = 1.12$, $p = .270$. Results are shown in Figure 3-5, panel *i*.

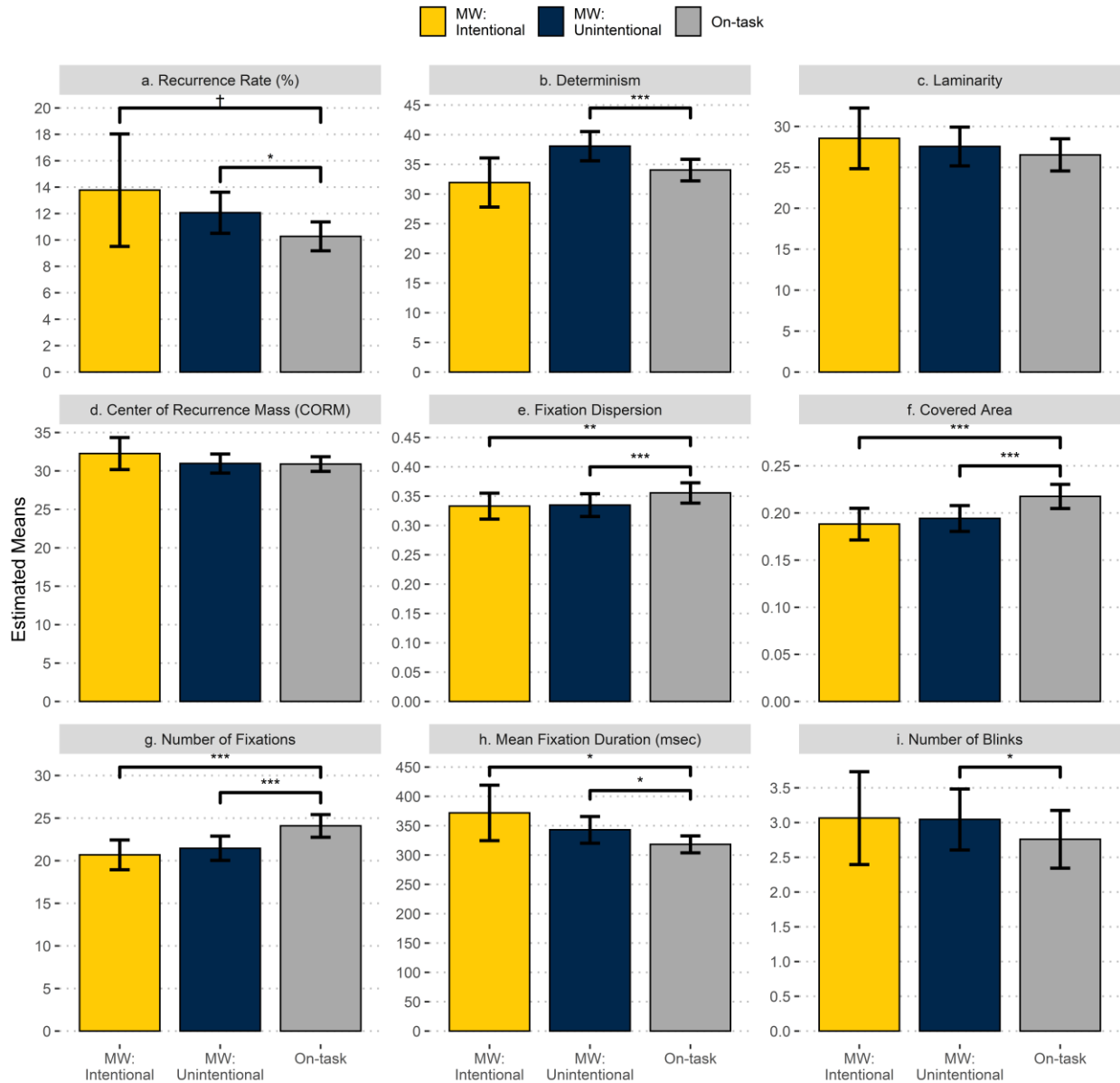


Figure 3-5. The estimated means of eye-tracking measures during intentional MW, unintentional MW, and on-task episodes. Panels *a* to *d* show measures from Recurrence Quantification Analysis. Panels *e* and *f* show measures indicating the spatial property of fixations. Panel *g* to *i* show additional measures examined. Results are aggregated across the stimuli type. Error bars show 95% confidence intervals. The annotations indicate the significance level of the difference between intentional/unintentional MW and the on-task episode. ** $p < .001$, * $p < .01$, * $p < .05$, † $p < .10$.

3.4.11 Relationship between meaning and saliency maps.

We calculated the linear correlation between the meaning maps and the saliency maps by correlating corresponding pixels for each picture. The results are shown in Figure 3-6. The average correlation between meaning maps and saliency maps was 0.85 ($SD = .06$).

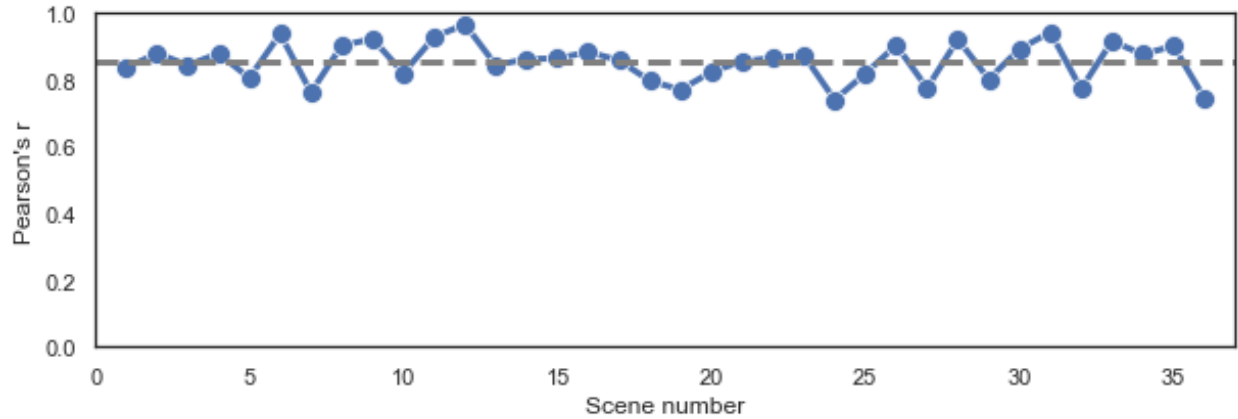


Figure 3-6. Correlation between meaning and saliency maps for each scene. The X-axis represents the scene number. The Y-axis represents the linear correlation between meaning and saliency maps. The horizontal dashed line represents the average value across all scenes.

3.4.12 Relationship between fixations and scene properties

Following Henderson and Hayes (2017), we used squared linear correlations to quantify the shared variance between attention maps and meaning/saliency maps. The mean values of these measures are shown in Figure 3-7.

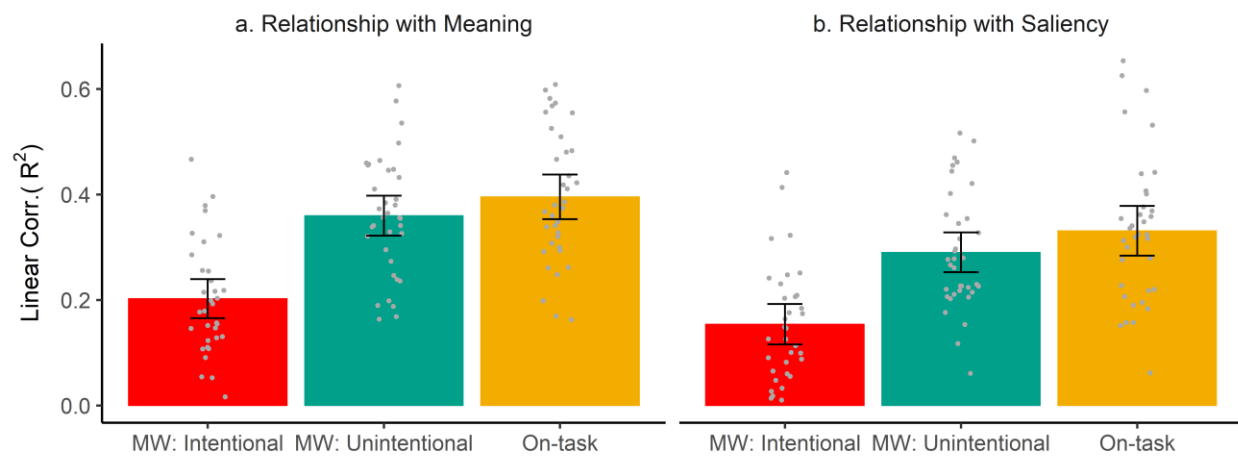


Figure 3-7. Squared linear correlations between image properties and fixation allocation during on-task and MW episodes. Panel *a* shows relationships with meaning maps and panel *b* shows relationships with saliency maps. Error bars show 95% confidence intervals. The gray dots represent raw data points.

A two-way ANOVA with attentional state (on-task, intentional MW, & unintentional MW) and map type (meaning, saliency) as within-item variables showed a significant main effect of attentional state, $F(1.77, 58.27) = 88.53, p < .001, \eta_G^2 = .33$. There was also a significant main effect of map type, $F(1, 33) = 12.65, p = .001, \eta_G^2 = .07$. The interaction between attentional state and map type was not significant, $F(1.52, 50.02) = 1.78, p = .19, \eta_G^2 = .002$.

We then conducted pairwise comparisons (Holm-Bonferroni corrected) to evaluate differences among attentional states separately for meaning and saliency. Fixations during on-task episodes were more strongly associated with meaning compared to those during unintentional MW, $b = .04, t(92.1) = 2.30, p = .02$, and those during intentional MW, $b = .20, t(92.1) = 12.11, p < .001$. There was also a significant difference between unintentional MW and intentional MW in their relations to meaning, $b = .16, t(92.1) = -9.80, p < .001$. Similarly, on-task fixations were more strongly associated with saliency than those from unintentional MW, $b = .04, t(92.1) = 2.47, p = .02$, and intentional MW, $b = .18, t(92.1) = 10.85, p < .001$. We also found a significant difference between unintentional MW and intentional MW in their relations to saliency, $b = .14, t(92.1) = -8.38, p < .001$.

3.4.13 Re-sampling on-task fixations.

In this study, MW reports were relatively infrequent compared to on-task reports. Attention maps produced from an insufficient number of fixations may be less correlated with scene properties, to begin with, compared to those produced from a large number of fixations. Thus, the observed differences between on-task and MW could simply be due to the difference in sample sizes. To examine this possibility, we re-sampled on-task fixations to match the total

number of fixations in the corresponding MW episodes. For example, if there were 100 intentional MW fixations for a given picture (across participants), we would correspondingly sample 100 fixations from the pool of on-task fixations. We then “drew” an attention map based on the reduced on-task sample and computed the squared linear correlation with meaning/saliency as in the previous section. A paired-sample *t*-test was used to compare values produced from the reduced on-task samples across pictures against the original MW data. This procedure was done 1000 times for comparing with intentional MW, and 1000 times for comparing with unintentional MW. The produced *t*-values were plotted as histograms in Figure 3-8.

As shown in Figure 3-8, panel *a* and *b*, there were more than 95% of the *t*-values greater than the critical *t* value (2.04), indicating that the differences between on-task and intentional MW were not simply due to difference in sample sizes. However, as shown in Figure 3-8, panel *c* and *d*, more than 95% of the *t* values were below the critical *t* for the comparisons involving unintentional MW. This finding casts doubt on whether unintentional MW truly differed from on-task episodes in terms of the association between fixations and scene properties.

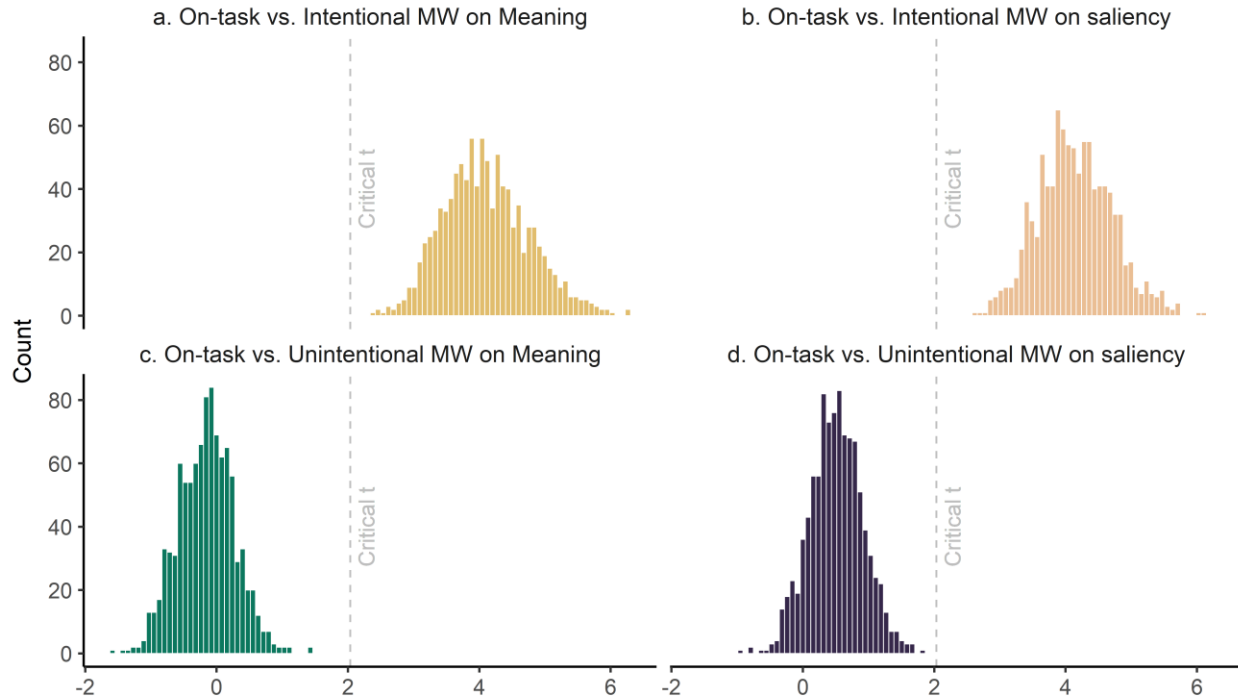


Figure 3-8. Histograms of t -values from paired-sample t -tests comparing R^2 values during MW to R^2 values calculated from subsets of on-task data with the same number of fixations. Panel *a* shows t -values comparing on-task and intentional MW on meaning. Panel *b* shows t -values comparing on-task and intentional MW on saliency. Panel *c* shows t -values comparing on-task and unintentional MW on meaning. Panel *d* shows t -values comparing on-task and unintentional MW on saliency. The vertical dashed line represents the critical t value in each case for declaring significance.

3.4.14 Meaning vs. saliency during MW.

The ANOVA results of the squared linear correlations suggest that fixations were more strongly associated with meaning than saliency regardless of attentional states. Because meaning and saliency were highly correlated, we further used squared semi-partial correlations to examine the ability of meaning and saliency to independently account for the variance in fixations in different attentional states. Results from a two-way repeated-measure ANOVA showed that the main effect of attention was not significant, $F(1.41, 46.63) = 2.33, p = .12, \eta_G^2 = .009$. However, there was a significant main effect of map type, $F(1, 33) = 12.64, p = .001, \eta_G^2 = .13$. The interaction between attention and map type was not significant, $F(1.51, 49.99) = 1.78, p = .19,$

$\eta_G^2 = .004$. We then conducted pairwise comparisons to evaluate the difference between meaning and saliency maps at each attentional state. Meaning maps, compared to saliency maps, accounted for more *unique* variance at all attentional states: on-task, $b = .07$, $t(46.4) = 3.56$, $p < .001$; unintentional MW, $b = .07$, $t(46.4) = 3.70$, $p < .001$; intentional MW, $b = .05$, $t(46.4) = 2.49$, $p = .02$.

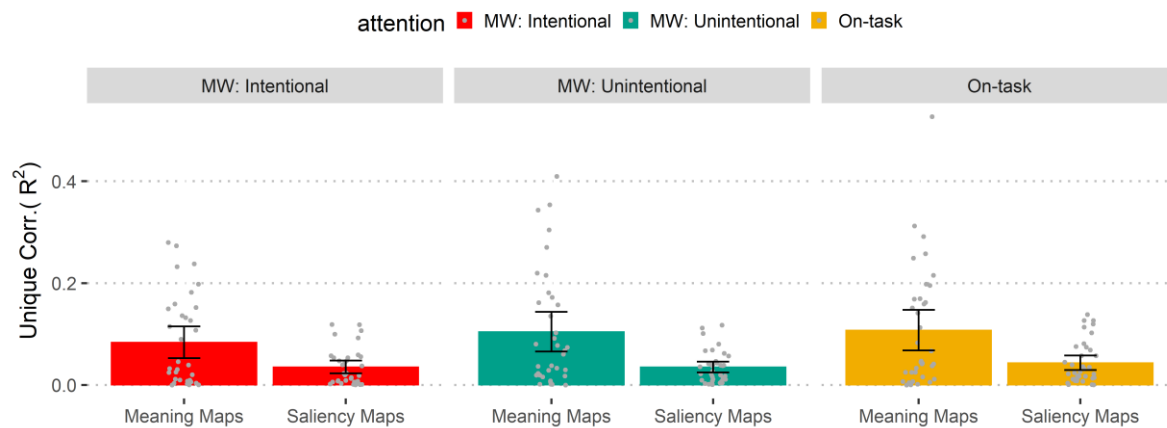


Figure 3-9. Squared semi-partial correlations of meaning and saliency maps in different attentional states. Error bars show 95% confidence intervals. The gray dots represent raw data points.

3.5 Discussion

The current study conducted a comprehensive examination of eye movements during MW in a scene perception task. We extended previous research by focusing on the sequential pattern (i.e., recurrence) of eye movements and examining how the relationship between attention and the scene changes during MW. As shown, MW was associated with significant changes in both aspects.

3.5.1 Recurrence patterns.

We found that scanpaths were more repetitive during unintentional MW than during on-task episodes, as indicated by increased recurrence and determinism. It seems as if participants during unintentional MW were re-enacting their previous visual processing by repeating their scanpaths. According to Zelinsky et al. (2011)'s "monitor-refixate-rehearse" model, an object's representation in visual working memory rapidly decays after the gaze moves to new regions and a monitoring system can direct the gaze back to "refresh" its fading representation. Object representations in visual working memory might undergo serious interference during unintentional MW, increasing the likelihood of refixations. Moreover, the increased determinism during unintentional MW indicates that a larger proportion of fixation sequences were repeated. Repeating the same scanpath is known to aid the rehearsal of visual-spatial information (Tremblay, Saint-Aubin, & Jalbert, 2006), which suggests that the rehearsing mechanism might go beyond single refixations. Overall, the repetitive scanpath patterns may reflect an adaptive response to failures in perception due to MW, with the ability to monitor memory interference and to execute returning saccades still functional. Participants might be able to remember where they looked at even during MW and use this information to guide their eye movements. This could involve relatively more automated processes compared to maintaining novel object representations in visual working memory and therefore be less subject to effects of MW. That said, having this monitoring mechanism alone does not seem enough since memory was worse for images where participants reported mind-wandering.

One might reason that the increased recurrence and determinism during unintentional MW were caused by participants repeatedly looking at the same location. Indeed, this viewing pattern would produce recurrence clusters on the recurrence plot that elevate both recurrence and

determinism. However, this explanation is complicated by the lack of difference in laminarity and CORM between on-task and unintentional MW. Laminarity is a direct measure of repeatedly fixating the same region but there was little evidence that there was increased laminarity during unintentional MW. Moreover, repeatedly looking at the same location also implies that refixations occur soon after the initial fixation, which would lead to a smaller CORM value. Again, we did not find this to be the case. Therefore, it does seem that participants were revisiting different previous locations in the same order during unintentional MW, resulting in a somewhat stereotypical pattern of looking.

It is also worth noting that determinism increased during unintentional MW but not during intentional MW, whereas the recurrence rate for both types of MW followed a similar pattern (although it was only significant for unintentional MW vs. on-task). Relatedly, intentional MW showed the highest CORM value (although it was non-significant), suggesting that the refixations occurred relatively late. One speculation is that, during intentional MW, participants deliberately engaged in off-task thoughts soon after the presentation of an image, and then came back to re-process the image when the presentation time was almost up. This attentional pattern can lead to low determinism because of the limited time to form long repeated scanpaths. But controlling when the refixations would occur seems less likely during unintentional MW due to its uncontrolled nature (Seli, Risko, Smilek, et al., 2016).

The increased recurrence during MW potentially contributed to changes in the overall spatial distribution of fixations. We found that fixations were located closer to each other and covered less of the visual content during both types of MW. These findings were not consistent with Krasich et al. (2018), in which they observed a larger fixation dispersion during MW. However, there are some important differences between the two tasks. First, a part of the

memory test in Krasich et al. (2018) required participants to identify small vignettes extracted from the full stimuli, which might have encouraged a skimming pattern during MW. Second, the stimuli in Krasich et al. (2018) contained a large number of human faces and other complex objects. If during MW, participants prefer to look at human faces over other stimuli (Zhang, Miller, Sun, & Cortina, 2020) and human faces happen to distribute widely across the scene, fixation dispersion may increase. Third, the stimulus presentation time (45 - 75 sec) in Krasich et al. (2018) was much longer compared to the current study (10 sec) and the increase in dispersion was only significant towards the end of presentation time. It is unknown whether we would have observed a similar pattern if we had extended the presentation time.

3.5.2 Previous measures.

We found that MW was associated with fewer and longer fixations, similar to those in Krasich et al. (2018). These results are broadly consistent with those observed during reading comprehension (Faber et al., 2018; Foulsham et al., 2013; Reichle et al., 2010; Steindorf & Rummel, 2019). This similarity suggests that there are core features of MW that are independent of the specific task environment. From the perspective of the perceptual decoupling account, the decreased fixation count may indicate an overall reduction of external processing, and the increased fixation duration may indicate the reduced efficiency at processing perceptual information. Alternatively, fewer and longer fixations may simply indicate low arousal which may occur during MW (Unsworth & Robison, 2018).

We also found that unintentional MW was associated with more blinks. A numerically similar pattern was observed for intentional MW. While the results are generally consistent with those in Smilek et al. (2010), the estimated differences ($b = 0.28$ for unintentional MW and $b =$

0.30 for intentional MW) seem to be considerably smaller than the original study. In our analysis, we set the maximum blink duration to be 500 ms. Signal loss beyond 500 ms can indicate micro-sleeps or a substantial change in the head position, both of whom are likely to occur when task engagement is low (Caffier, Erdmann, & Ullsperger, 2003; Häkkinen et al., 1999; Schleicher, Galley, Briest, & Galley, 2008; Seli et al., 2014). As an additional analysis, we analyzed the frequency of “blinking” episodes that exceeded 500 ms. We found that these incidents indeed occurred more often during unintentional MW than during on-task episodes, $b = 0.18$, $SE = 0.05$, $t = 3.43$, $p = .001$. A similar pattern was observed for intentional MW but the difference was not significant, $b = 0.17$, $SE = 0.14$, $t = 1.24$, $p = .222$. Therefore, besides blinking, other forms of data loss may also occur more often during MW. Future research is clearly needed to further examine this issue.

3.5.3 Meaning and Saliency.

We initially found that both intentional and unintentional MW showed a reduced association between fixation allocation and scene meaning. However, an alternative explanation was that this was due to differences in the number of fixations available to generate the attention maps. To test this possibility, we equated the number of fixations by randomly sampling a subset of the on-task fixations. After applying this procedure, only intentional MW still showed the same pattern. Thus, we conclude that there was a robust difference between fixation allocation during on-task vs. during intentional MW in terms of its relationship with meaning. This finding provides some insights regarding the nature of the attentional decoupling process during MW, such that attentional decoupling may depend on the specific form of MW engaged by the participants.

The high correlation between meaning and saliency was consistent with previous research (e.g., Henderson & Hayes, 2017). Therefore, it is no surprise that the relationship with saliency followed a similar pattern with meaning. Importantly, we found that meaning maps still accounted for more *unique* variance than saliency maps did even during intentional and unintentional MW. The case of intentional MW is especially interesting because participants deliberately disengaged from the task. Thus, gaze control during real-world scene perception seems to be consistently based on meaning rather than saliency, even when meaning was irrelevant (Peacock et al., 2019). We reason that this meaning-based process might be highly automatic and being inattentive does not overhaul this mechanism.

3.5.4 Putting together

So far, we have examined a large number of measures that capture the global, local, and sequential aspects of eye movements. How do these results fit together? First, both intentional and unintentional MW showed a numerically higher recurrence rate, although in addition to this, the recurrent fixations were also more deterministic during unintentional MW. Second, the high recurrence rate potentially affected how fixations are distributed spatially – for both unintentional and intentional MW, fixations were located closer to each other and covered less area of the stimuli. Third, there were fewer fixations during MW compared to being on-task. Taken together, these results suggest an eye movement pattern in which fixations are limited in number and recurrence in space. Participants with such eye movement patterns may overlook certain meaningful regions of the scene, leading to a reduced association between attention and the meaning map.

It is interesting to note that the relationship between attention and meaning seems to be less affected during unintentional MW (compared to intentional MW), even though both types of MW demonstrated similar global characteristics. In this study, the definition of unintentional MW was communicated to the participants as “your thoughts drifted away *despite your best intentions to focus on the task*”, whereas intentional MW as “you intentionally decided to think about things that are unrelated to the task”. Given participants’ original intention to stay on task during unintentional MW, they might look at things similar to what they would look if they were on-task. On the other hand, participants during intentional MW may look at more meaningless things. One previous study (Phillips et al., 2016) found that, when given a chance to re-read an article, participants read at a much faster pace compared to first-pass and also mind-wandered more often. Critically, the re-reading effect of MW was completely driven by intentional MW. Thus, it appears that intentional MW is associated with substantial changes in attention and eye movements. On the other hand, the present results are inconsistent with the idea that people can to some extent intentionally engage in off-task thoughts without affecting primary task performance (e.g., Seli, Carriere, et al., 2018). Both intentional and unintentional forms of MW were associated with worse memory of the target pictures.

3.5.5 Limitations and future directions.

A common challenge for MW research is that there are often no effective ways to externally manipulate MW. As a result, studies have typically employed the thought-probe method to obtain MW data. While responses to the thought probes do seem to be valid indications of participants’ internal state (Smallwood & Schooler, 2015), this procedure often produces heavily unbalanced data. The overall MW rate in our study ($Mean = 27\%$, $SD = 22\%$) was similar to a previous study examining MW during scene perception ($Mean = 27\%$, $SD =$

22% in Krasich et al., 2018). But Krasich et al. (2018) did not further differentiate between intentional and unintentional MW. In our study, the rate of intentional MW was quite low. To our knowledge, this is one of the first studies examining eye movement patterns during intentional MW. MW was found less likely to occur when there are rapid changes in external events (Faber et al., 2018). In our case, the serial presentation of novel and meaningful stimuli may somehow reduce MW and especially intentional MW. While this was undesirable for our study's purpose, it may be further explored as an intervention method to reduce MW. Our preliminary results suggest that intentional and unintentional MW showed similar patterns for fixation count, duration, dispersion, and eye blinks. But differences might exist in terms of how eye movements unfold over time and how they are related to the properties of the scene. Because MW is a wide-ranging term, different sub-types of MW might not share the same set of gaze signatures.

It is also worth noting that RQA only captures a single aspect (i.e., recurrence) of the highly complicated information existing in scanpaths, and there are many other analysis methods available. For example, the MultiMatch technique (Dewhurst et al., 2012) captures the similarity between two scanpaths in various dimensions, such as shape, direction, length, and duration. ScanMatch (Cristino, Mathôt, Theeuwes, & Gilchrist, 2010), as another example, computes the overall similarity between scanpaths. These methods provide alternative ways to analyze eye movements during MW and on-task episodes. For example, scanpaths might be more similar to each other when attention is coupled to the task but become more idiosyncratic when individuals are mind-wandering. In sum, we believe that scanpath analysis, as a whole, provides a unique perspective on the current literature of MW and bears the potential to uncover more information about its underlying mechanisms.

There is an ongoing effort to develop gaze-based detectors of MW, but so far the research has been primarily focusing on global measures, which do not consider fixation sequence or the content of the stimuli (Bixler & D’Mello, 2016; Faber et al., 2018; Hutt et al., 2017). We believe that the current results will be helpful not only for understanding the mechanisms of attention but also for diagnosing moments of inattention. Given the fundamental role of scene perception in daily activities (e.g., driving), knowing when people are mind-wandering can have important implications for public welfare. We hope that our work will provide information useful to this ongoing line of research and lead to more studies on how the vagaries of attention and the features of stimuli and tasks combine to determine how looking unfolds over time.

Chapter 4 Mind-wandering during Reading

4.1 Abstract

The current study examined whether mind-wandering was associated with reduced re-reading when the reader reads so-called "garden-path jokes." In a garden-path joke, the reader's initial interpretation is violated by the final punchline, and the violation creates a semantic incongruity that needs to be resolved (e.g., "My girlfriend has read so many negative things about smoking. Therefore, she decided to quit reading."). Rereading text before the punchline can help resolve the incongruity. Participants read jokes and non-funny controls embedded in filler texts and responded to thought probes that assessed intentional and unintentional mind-wandering. When the reader was not mind-wandering, jokes elicited more rereading (from the punchline) than the non-funny controls did, and had a recall advantage over the non-funny controls. During mind-wandering, however, the additional eye movement processing and the recall advantage of jokes were generally reduced. These findings show that mind-wandering hinders rereading, which is important for resolving higher-level comprehension difficulties.

4.2 Introduction

Reading comprehension is susceptible to mind-wandering, a mental state in which attention shifts from the external task to self-generated, task-irrelevant thoughts (Smallwood & Schooler, 2015). How does mind-wandering change the way people read? Can these changes reveal impairments of the cognitive processes underlying reading? During the past few years, an increasing number of studies have used eye-tracking to study these questions (Faber et al., 2018; Foulsham et al., 2013; Reichle et al., 2010; Schad, Nuthmann, & Engbert, 2012; Uzzaman &

Joordens, 2011). One benefit of using eye-tracking is its direct examination of the “eye-mind” link—the extent to which cognition actively controls what people are looking at. But, due to reduced top-down control of comprehension, this link may break down during mind-wandering.

The normal reading process can be generally described as going through a hierarchy of stages, from extracting lexical meanings from printed words (Pollatsek et al., 2006), to integrating words into propositions (Frazier, 1998), and finally to establishing a coherent understanding of the entire passage (Zwaan & Radvansky, 1998). Several eye-tracking studies have shown that during mind-wandering, the normal association between fixation duration and lexical properties of the word (e.g., longer looking times for low-frequency words) was reduced (Foulsham et al., 2013; Reichle et al., 2010; Schad et al., 2012), suggesting deficits during lexical processing.

Smallwood (2011) reasoned that impairments in the early stages of reading can have implications for later processes, so mind-wandering should have profound impacts on higher-level processes. Extant studies examining higher-level processes have typically used self-paced reading (for an exception, see Schad et al., 2012). One study asked participants to read “gibberish” texts that changed the order of nouns or pronouns (as described in Smallwood, 2011). Not being able to detect gibberish texts quickly, according to the authors, would indicate impairment in the creation of propositions. Results showed that when readers were mind-wandering, they were likely to keep reading without noticing that the text had become gibberish. Another study (Smallwood, McSpadden, & Schooler, 2008) asked participants to read a Sherlock Holmes story word-by-word and found that, if participants were mind-wandering when critical clues about the villain were offered, they were less likely to correctly infer the identity of the

villain. The authors argued that mind-wandering at critical points interfered with the integration of important events necessary to identify the villain.

In the self-paced reading paradigm, participants can see only one word at a time and are not permitted to look back at previous portions of the text. However, during free reading, about 10% to 15% of saccades move backward to previous text (Rayner, 1998). One important reason for making such regressions is to resolve difficulties during higher-level stages of comprehension (for a review, see Bicknell & Levy, 2011). Therefore, studying how rereading behavior is affected during mind-wandering can advance our understanding of the mental state's effect on reading. Interestingly, previous studies did not find consistent evidence that rereading was affected during mind-wandering (for a review, see Steindorf & Rummel, 2019), possibly because participants were not processing texts in which rereading is critical for comprehension.

In what situation do people tend to reread? One example is when they read the so-called garden-path jokes (Dynel, 2009). Garden-path jokes elicit humor by violating the reader's original interpretation of the text at the final punchline. To "get" the joke, the reader must resolve the semantic incongruity, or in other words, find a new interpretation of the text (Suls, 1972, 1983), for example, "*For more than 40 years, I have only loved one woman. I hope my wife will never know*" (Mayerhofer & Schacht, 2015).

In a garden-path joke, the set-up is designed to be compatible with at least two interpretations. However, to the reader, one interpretation is highly salient, as determined by the reader's general world knowledge. Thus, the reader is "tricked" to adopt the salient interpretation before encountering the punchline. In the previous example, readers may wrongly assume that the set-up describes a loyal husband. However, this interpretation is violated at the punchline, causing a semantic incongruity. Thus, the reader must backtrack the set-up to search for the

covert interpretation to resolve the difficulty. For example, the reader may adopt a new interpretation that the husband has been cheating on his wife for 40 years. The successful resolution of semantic incongruity allows for a sense of amusement (Dyner, 2009). Note that a non-funny but coherent version of the joke can be constructed by simply replacing “know” to “forget.” Doing so will reduce the text’s semantic incongruity and humor potential. The incongruity-resolution theory (Suls, 1972, 1983) and its variations (Coulson & Kutas, 1998; Mayerhofer & Schacht, 2015; Ritchie, 2004) constitute a well-established framework that describes the cognitive processes of humor processing (Dyner, 2009; Mayerhofer & Schacht, 2015).

The incongruity-resolution process of garden-path jokes can be indexed by behavioral and physiological measures. In a self-paced reading task, joke endings received longer reading time than the ending of nonfunny control sentences did (Mayerhofer & Schacht, 2015, Experiment 1). Electroencephalography data showed that joke endings elicited a larger N400 component compared to coherent endings, indicating semantic integration difficulties (Mayerhofer & Schacht, 2015, Experiment 2 and 3). Importantly, Coulson, Urbach, and Kutas (2006) used a free reading paradigm and showed that garden-path jokes, compared to nonfunny controls, produced more rereading eye movements from the ending. This finding, according to the authors, shows a processing cost due to the construction of an alternative cognitive model of the text (Coulson et al., 2006).

Some important features distinguish garden-path jokes from traditional garden-path sentences (e.g., “The horse raced past the barn fell”; Frazier & Rayner, 1982) and gibberish texts. The incongruity and its resolution of garden-path jokes are localized at the semantic level rather than the syntactic level. In other words, the reader is prompted to discover an alternative

meaning rather than an alternative parsing. Other researchers have described this process as a frame-shifting (Coulson & Kutas, 1998), a forced reinterpretation (Ritchie, 2004), or a belief revision (Mayerhofer & Schacht, 2015), all of which point to a reanalysis at the semantic level. During this process, the reader must consult their general world knowledge or previous experience to reinterpret the linguistic input. Thus, the resolution of comprehension difficulties occurs at an advanced level of understanding and requires a close coupling between attention and linguistic input. This may make its processing highly susceptible to mind-wandering (Schad et al., 2012). Moreover, compared to gibberish texts, garden-path jokes are intelligible, which might render them more ecologically valid. In sum, we believe that garden-path jokes provide a promising opportunity to study how mind-wandering affects higher-level processes of reading.

4.2.1 The current study.

The current study sought to investigate whether mind-wandering affected the resolution of semantic incongruity, a higher-level cognitive process required for understanding garden-path jokes. Previous research has suggested that a critical index of this process is rereading from the punchline. Therefore, we recorded participants' eye movements while they read garden-path jokes and nonfunny controls embedded in filler texts. Participants responded to thought probes after each joke and control text to report mind-wandering. Our hypothesis was straightforward: The incongruity-resolution process was present when attention was on the task but was impaired during mind-wandering.

Mind-wandering encompasses a wide range of mental experiences that vary in numerous dimensions (Seli, Kane, Smallwood, et al., 2018). Recent evidence suggests that mind-wandering can emerge with or without intention (Seli, Risko, & Smilek, 2016). Unintentional mind-wandering reflects a spontaneous shift from task-related to task-unrelated thoughts, despite the

individual's willingness to stay on task. However, it is estimated that more than one-third of mind-wandering thoughts emerge with intention, a controlled and deliberate disengagement (Seli, Cheyne, et al., 2015; Seli, Wammes, Risko, et al., 2016). Previous research has shown that intentional and unintentional mind-wandering are sometimes dissociable. For example, increasing task difficulty reduces the rate of intentional mind-wandering but increases the rate of unintentional mind-wandering (Seli, Risko, & Smilek, 2016); task motivation correlates more strongly with intentional mind-wandering than with unintentional mind-wandering (Seli, Cheyne, et al., 2015). That said, both types of mind-wandering were found to impair task performance in a sustained attention task (Seli, Cheyne, et al., 2015) and a video lecture task (Seli, Wammes, Risko, et al., 2016). Their similar effects are not surprising, because both types of mind-wandering involve a decoupling of attention from the task at hand. In sum, it is important to treat mind-wandering not as a unitary concept, even if we predict that intentional and unintentional mind-wandering have similar effects on the incongruity-resolution process.

4.3 Methods

4.3.1 Participants.

Forty-seven undergraduate students from the University of Michigan (mean age = 18.96, $SD = .95$, 25 females) participated in the study for course credit. All participants were native English speakers with normal eyesight. Due to technical failures, three participants only completed half of the experiment. However, their data were included in analyses.

4.3.2 Apparatus and stimuli.

We obtained forty-six garden-path jokes and their corresponding non-funny control texts. Each joke-control pair shares the same texts until the ending. The jokes' endings were designed

to elicit humor by violating the previous set-up. The non-funny controls' endings were designed to be coherent and neutral. Thirty-nine joke-control pairs differ only in the final word, and the other seven pairs differ in the last two words. Besides, 480 neutral fillers were constructed to mimic the linguistic style (e.g., length, topic, difficulty, etc.) of the target sentences. Some examples are shown below. See the online supplemental material for full stimuli.

1. Joke: For more than forty years I have only loved one woman. I hope my wife will never know.

2. Control: For more than forty years I have only loved one woman. I hope my wife will never forget.

3. Filler: I walked into the grocery store. I was going there to buy my favorite energy drink.

As a manipulation check, we recruited sixty Mturk workers to rate the jokes and the non-funny controls on three scales: comprehensibility, funniness, and predictability of the ending. Each scale included three items. All items used a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). These rating scales were developed by Mayerhofer & Schacht, 2015 and were used to evaluate the garden-path jokes used in their study. Every worker rated twenty-three jokes and twenty-three controls. One joke-control pair had very low comprehensibility (2.58, the rest: Mean = 3.96, *SD* = 1.03). We included this pair in the experiment for the convenience of constructing stimuli presentation orders (as described later in this section), but the data from this pair was discarded from all subsequent data analyses.

We used linear mixed models (lme4; Bates et al., 2014) to examine differences between the jokes and the non-funny controls with maximum random effects. We used the lmerTest package (Kuznetsova et al., 2017) to obtain an approximation of *p*-value. Results showed that,

compared to controls, jokes were rated as funnier ($b = 1.05$, $SE = 0.12$, $t = 9.10$, $p < .001$) and had less predictable endings ($b = -0.76$, $SE = 0.10$, $t = -7.80$, $p < .001$). However, the jokes were not significantly less comprehensible than controls ($b = 0.12$, $SE = 0.07$, $t = 1.71$, $p = 0.09$). In Figure 4-1, we showed mean ratings of jokes and controls on the three scales. In sum, the Mturkers' ratings confirmed the validity of our stimuli.

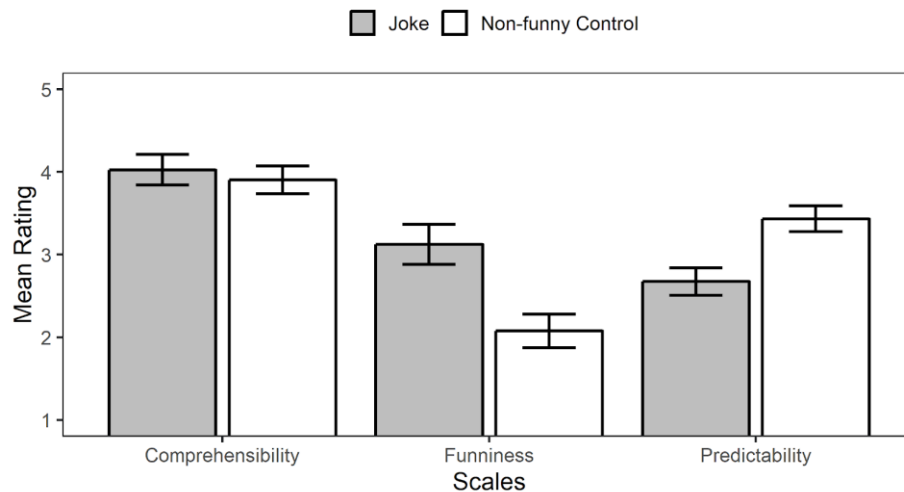


Figure 4-1. Mean Ratings of jokes and non-funny control sentences by Mturk workers. Error bars show the 95% Confidence Intervals.

Based on the forty-six joke-control pairs, we constructed sixteen pseudo-random stimuli presentation orders. In every order, (1) twenty-three of the texts appeared as jokes and the other twenty-three appeared as non-funny controls, and (2) each joke and non-funny control was preceded by 5 to 15 fillers. We spaced out target texts with fillers to increase the distance between probes (the thought probe occurred after every joke and control), as frequently probing the participant can reduce mind-wandering reports (Seli, Carriere, Levene, & Smilek, 2013). The average distance between any two targets was ten fillers. This resulted in each participant reading 526 texts throughout the experiment: 46 target texts (23 jokes and 23 controls) embedded in 480 filler trials. We divided the whole experiment into two blocks of the same size: Both

blocks have 23 target trials (11 jokes and 12 controls, or vice versa) embedded in 240 filler trials.

Stimuli were presented on a 20.1-inch computer screen at approximately 70 centimeters to the participant. Text font was Times New Roman and text size was 37.5. Each letter subtended horizontally about 0.65 degrees of visual angle. Monocular eye movements were recorded by the Eyelink Remote System at a sampling rate of 500 Hz. To ensure comfort, no chin rest was used and head movement was adjusted by tracking a sticker on participants' forehead. The experiment was implemented using the OpenSesame software (Mathôt et al., 2012) with functions from the PyGaze package (Dalmaijer et al., 2014).

4.3.3 Procedure.

The experiment began with a survey asking all participants to make a to-do list for the next five days, as future-planning has been shown to increase the mind-wandering rate during a subsequent task (Seli, Cheyne, et al., 2015). Then, we introduced the reading task to the participants. Participants were asked to read sentences for comprehension. Participants at this point did not know the existence of jokes or the nature of the test afterward. This was done to eliminate the possibility that re-reading results from participants purposefully memorizing the jokes rather than the incongruity-resolution process per se. Next, participants were told that, during reading, a thought probe will occur occasionally, which requires them to report whether they were “on-task” or “off-task” during the previous text. The experimenter introduced the definitions of “on-task” and “off-task”: *Being on-task means that, just before the screen appeared, you were focused on completing the task and were not thinking about anything unrelated to the task.* “Off-task” means that just before the screen appeared, you were thinking about something completely unrelated to the task (Seli et al., 2015). Because the framing of

thought probes can affect reported mind-wandering rates (Weinstein, 2018), we used a neutral question (“*Just now where was your attention?*”). Participants were asked to answer “on-task” or “off-task” by pressing the corresponding key. We also randomly switched the order of “on-task” and “off-task” options across participants to reduce any confounds due to ordering. If “off-task” was chosen, participants were further asked to indicate whether mind-wandering was intentional or unintentional. Intentional mind-wandering was defined as “you intentionally decided to think about things that are unrelated to the task”, and unintentional mind-wandering was defined as “your thoughts drifted away despite your best intentions to focus on the task” (Seli et al., 2015). The order of this question’s options was also randomized across participants. We assigned participants to one of the sixteen stimuli orders based on their participant number. After calibrating the eye tracker, participants completed five practice trials. Each trial started with a fixation dot located at the position of the first letter of the upcoming text. The text appeared once a stable gaze signal at the dot was detected. Together with the text, there was also a small fixation dot at the bottom-right corner of the screen. Participants were asked to move to this dot once they have finished this trial. The trial ended once a stable gaze signal was detected at this dot. Then, after calibrating the eye tracker again, the experimental trials started. The task proceeded in an automated fashion. The thought probe occurred after every target sentence (i.e., jokes/non-funny controls). A research assistant quietly sat outside the participant’s field of vision and monitored the gaze-overlaid stimuli on a second monitor. Re-calibration was conducted if tracking quality deteriorated. After reading, participants were asked to complete a recall test to fill out the ending of each target trial (i.e., the part that was different between jokes and controls) with the previous text given. There was no time limit for this test. The entire experiment took about 120 minutes to finish.

4.3.4 Data analysis.

Fixations that were greater than 1,500 ms or shorter than 80 ms were discarded (3.92% of data). We chose a relatively high upper bound because mind-wandering was known to produce longer fixation duration compared to normal reading (Faber et al., 2018; Reichle et al., 2010). Because the incongruity-resolution process strictly speaks to what happens after the reader encounters the punchline, the analysis region was set to where the jokes and controls differ. In the previous example, the analysis region would be the word “know” for jokes and “forget” for controls. For the seven joke-control pairs that differed in the last two words, the analysis region included both words.

We examined the following measures: (a) recall (a binary variable indicating whether the answer matches the original text), (b) regressions– out (a count variable indicating the number of regressions from the analysis region to previous words), (c) regression– path duration (the sum of all fixations from entering the analysis region to the last fixation on the entire text), (d) gaze duration (the sum of all fixations from entering the analysis region for the first time until leaving the region), (e) total looking time (the sum of all fixations on the analysis region), and (f) skipping (a binary variable indicating whether the analysis region was not fixated on throughout the trial). We used recall performance as an offline measure of the incongruity-resolution process. If jokes received additional visual processing (compared to controls), we expected that this should translate to better memory of the endings (Strick, Holland, van Baaren, & Van Knippenberg, 2009). Thus, we expected a significant recall advantage for jokes (compared to controls) only when participants indicated being on-task. Regressions-out and regression-path duration are critical measures for this study because they can indicate the degree to which participants reanalyzed the text from the ending. We expected more such rereading for jokes than

for controls, but only when the reader was on-task. We used gaze duration, total looking time, and skipping as supplemental measures. They do not directly speak to the rereading process but offer important details of how the ending was processed. Gaze duration, in relation to total looking time, measures early stages of language processing because it only includes first-pass reading. Coulson et al. (2006) found that gaze duration was not statistically different between jokes and controls, but they found a trend for longer total looking time for joke endings. We included these two measures to compare our results to previous research. Finally, not skipping the ending is likely a prerequisite for the incongruity-resolution process. A joke's ending might be less likely to be skipped than a nonfunny control's ending, but this effect, if true, should only occur when the participant was on-task.

We conducted a set of a priori contrasts to analyze the measures (Ruxton & Beauchamp, 2008; Schad, Hohenstein, Vasishth, & Kliegl, 2018). We created four orthogonal contrasts: one contrast for the effect of text type (joke/control) for each type of attention (on-task/ intentional mind-wandering/unintentional mind-wandering), and an additional contrast for the difference between mind-wandering and non-mind-wandering conditions. The fourth contrast was exploratory and tested how sentence endings, aggregating over jokes and controls, were processed during mind-wandering and non-mind-wandering. A weight matrix for the contrasts can be found in the online supplemental material. A regression model was built for each of the six dependent measures. Duration measures were log-transformed to fit to linear mixed models. Binary and count measures were modeled by generalized linear mixed models (GLMMs). Specifically, recall and skipping were modeled by binomial GLMMs with a logit link. Regressions-out were modeled by a Poisson GLMM with a log link (the default option). For convenience and clarity, in all models, we collapsed text type and attention into a single variable

of six groups called condition. We applied our custom contrasts to *condition*. Because word length and word frequency were known to influence eye movements (Kliegl et al., 2004; Rayner, 1998), and because the jokes and controls were not equated on these measures, we included word length and the logarithm of word frequency as covariates in all models of eye movement measures.¹ Random effects included (a) variations across participants, (b) variations across text frames, (c) variations for each (observed) combination of participant and condition, and (d) variations for each (observed) combination of text frames and condition. The R package *lme4* (Bates et al., 2014) was used for all model-fitting. Approximations of *p* values came from the *lmerTest* package (Kuznetsova et al., 2017).

4.4 Results

Overall, we obtained 1,195 on-task trials (58.35%), 546 unintentional mind-wandering trials (26.66%), and 307 intentional mind-wandering trials (14.99%). Additional details about the number of trials in each condition for each measure can be found in the supplemental materials.

4.4.1 Recall performance.

The probability of correct answers in each condition is shown in panel *a* of Figure 4-2. When participants indicated they were on-task, joke endings were more likely correctly recalled than neutral endings were, $b = 0.59$, $SE = 0.22$, $z = 2.74$, $p = 0.01$. However, this recall advantage was reduced during unintentional mind-wandering, $b = 0.55$, $SE = 0.29$, $z = 1.91$, $p = 0.06$, and was eliminated during intentional mind-wandering, $b = -0.04$, $SE = 0.39$, $z = -0.11$, $p = 0.91$. For the fourth contrast, recall was better when participants were on-task compared to when they were mind-wandering, $b = 1.16$, $SE = 0.17$, $z = 7.00$, $p < .001$.

¹ For the seven pairs that differed in the last two words, we used their total length and frequency of the phrase (from the Corpus of Contemporary American English; Davies, 2008). Results were similar without these covariates.

4.4.2 Eye movement measures.

Two critical indices of incongruity resolution were regressions-out and regression-path duration. Their marginal means were shown in panels *b* and *c* of Figure 4-2, respectively. When participants were on-task, jokes, compared to non-funny controls, elicited more regressions-out, $b = 0.28$, $SE = 0.11$, $z = 2.56$, $p = 0.01$. However, this difference was not significant during either unintentional mind-wandering, $b = -0.15$, $SE = 0.19$, $z = -0.80$, $p = 0.42$, or intentional mind-wandering, $b = 0.11$, $SE = 0.26$, $z = 0.43$, $p = 0.67$. For the last contrast, participants produced more regressions-out in general when they were on-task than when they were mind-wandering, $b = 0.38$, $SE = 0.10$, $z = 3.77$, $p < .001$.

Similarly, regression-path duration was longer for jokes than for controls when participants were on-task, $b = 0.09$, $SE = 0.03$, $t = 2.74$, $p = 0.01$. But this difference was not significant during either unintentional mind-wandering, $b = -0.02$, $SE = 0.05$, $t = -0.36$, $p = 0.72$, or intentional mind-wandering, $b = 0.08$, $SE = 0.07$, $t = 1.18$, $p = 0.24$. Finally, an overall difference was observed between on-task and mind-wandering, $b = 0.08$, $SE = 0.03$, $t = 2.84$, $p = .005$.

We then looked at gaze duration (Figure 4-2, panel *d*) and total looking time (Figure 4-2, panel *e*) on the analysis region. For gaze duration, we did not find a significant difference between jokes and controls even when participants indicated being on-task, $b = 0.02$, $SE = 0.02$, $t = 1.26$, $p = 0.21$. The difference was also not significant during unintentional mind-wandering, $b = .002$, $SE = 0.02$, $t = 0.36$, $p = 0.72$, or intentional mind-wandering, $b = 0.04$, $SE = 0.03$, $t = 1.26$, $p = 0.21$. There was also no significant difference in gaze duration between on-task and mind-wandering in general, $b = -.002$, $SE = 0.01$, $t = 0.01$, $p = 0.99$.

On the other hand, jokes produced significantly longer total looking time than controls did, when participants were on-task, $b = 0.04$, $SE = 0.02$, $t = 2.25$, $p = 0.02$. However, there was no significant difference during unintentional mind-wandering, $b = 0.00$, $SE = 0.03$, $t = -0.13$, $p = 0.90$, or intentional mind-wandering, $b = 0.07$, $SE = 0.04$, $t = 1.79$, $p = 0.07$. Total looking time did not significantly differ between on-task and mind-wandering in general, $b = 0.01$, $SE = 0.01$, $t = 0.98$, $p = 0.33$.

Finally, we looked at the probability of skipping the analysis region (Figure 4-2, panel *e*). When participants were on-task, joke endings were no less likely to be skipped than control endings were, $b = -0.19$, $SE = 0.15$, $z = -1.24$, $p = 0.21$. Moreover, the difference between jokes and controls was not significant during unintentional mind-wandering, $b = -0.08$, $SE = 0.21$, $z = -0.36$, $p = 0.72$, or intentional mind-wandering, $b = 0.37$, $SE = 0.28$, $z = 1.32$, $p = 0.19$. However, there was less skipping overall when participants were on-task than when they were mind-wandering, $b = -0.33$, $SE = 0.12$, $z = -2.67$, $p = .008$.

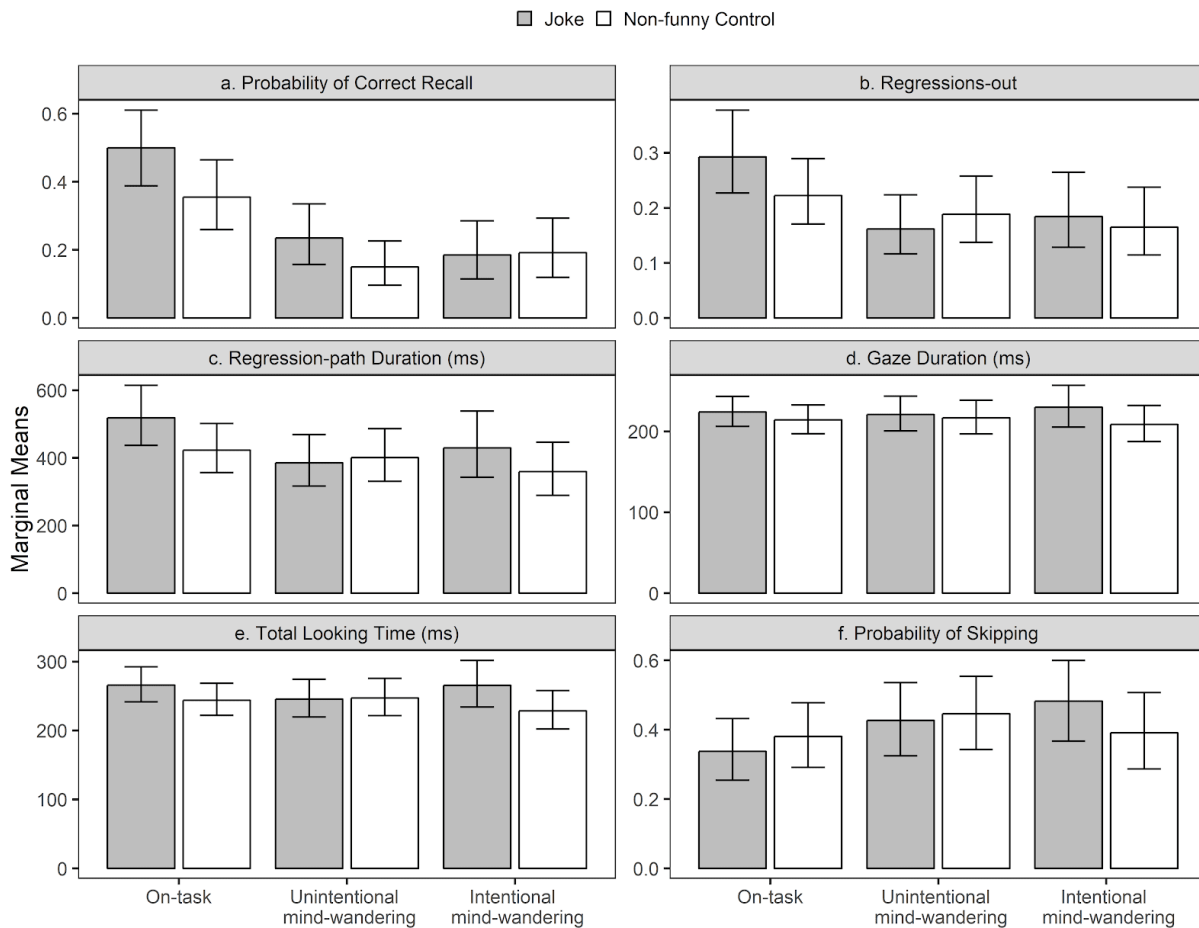


Figure 4-2. (a) Recall performance and (panel *b–f*) eye movement measures, by attention (on-task, unintentional mind-wandering, intentional mind-wandering) and text type (joke, control). Error bars show the 95% confidence intervals. All measures were back-transformed to the original scale. Eye movement measures were adjusted for word length and log word frequency.

4.5 Discussion

We examined how mind-wandering affected the semantic incongruity-resolution process of garden-path jokes. We hypothesized that the incongruity-resolution process would be impaired during both intentional and unintentional mind-wandering, but not when participants were on-task. The most important measures of this process were regressions– out and regression–path duration from the punchline. Our results show that, when participants were on-task, joke endings elicited more regressions– out and longer regression–path duration than

nonfunny controls did. These results provide a benchmark for how jokes (compared to controls) were processed without mind-wandering, which replicated Coulson et al. (2006)'s findings. However, the additional rereading of jokes was reduced during both intentional and unintentional mind-wandering, indicating impairments in the incongruity-resolution process.

We also examined several supplemental measures, including gaze duration, total looking time, and skipping. Similar to results in Coulson et al. (2006), only total looking time had a significant difference between jokes and controls when participants were on-task. Therefore, in addition to rereading previous texts, participants examined the punchline more than once, suggesting efforts of integrating the punchline and the set-up. This difference in total looking time was not observed during unintentional mind-wandering. Interestingly, for both gaze duration and total looking time, the intentional mind-wandering condition seemed to have a larger effect than the on-task condition did (although the differences were not significant in both cases). Perhaps during intentional mind-wandering, participants could sometimes notice the incongruity, leading to longer looking time at the ending. However, they did not put enough effort into rereading, presumably because of a lack of motivation.

For skipping, we did not find a significant difference in either the on-task, the intentional mind-wandering, or the unintentional mind-wandering condition. This finding is similar to that for gaze duration, as both speak to relatively early stages of reading. These findings suggest that the resolution of incongruity occurred at a relatively late stage, and it might not have been salient enough to affect early measures. Moreover, sentence endings naturally define processing units, and they might be important to look at for the control sentences as well.

Finally, we used recall performance as an offline measure of the incongruity-resolution process. If joke endings attracted additional attention, this would be reflected by how well

participants remembered the endings (Strick et al., 2009). Our results show that the recall advantage observed when participants were on-task was reduced during mind-wandering, which was consistent with the eye-tracking results. Importantly, this measure does not directly speak to whether participants really “got” the joke, a point we shall return to in the General Discussion.

Overall, our results show a clear pattern of how mind-wandering affected rereading and recall of garden-path jokes, signaling impairments in the incongruity-resolution process.

Following Experiment 1, we conducted a preregistered replication, to see if our major findings can be replicated.

4.6 Pre-registered Replication

We made some minor changes in the stimuli and procedure of Experiment 1, as specified in the sections below. All changes were preregistered. The preregistration protocol is available at <https://osf.io/jg27v/>.

4.6.1 Method.

Unless stated otherwise, the methodology remained the same as that in Experiment 1.

4.6.1.1 Participants.

We recruited 46 undergraduate students from the University of Michigan to participate in the study for course credit. According to the preregistered data exclusion criteria, we discarded data from three participants for technical failures, and three participants for not completing the entire experiment. The final sample size was 40 (Mage 18.85, SD .89, 23 female), which was specified in the preregistration. All participants were native English speakers with normal eyesight.

4.6.1.2 Stimuli.

In Experiment 1, one joke-control pair was rated to have low comprehensibility and seven joke-control pairs differed in the last two words. In the replication, we replaced them with eight new joke-control pairs that differed in only the last word. We recruited another 120 online workers to rate the new texts on the same scales used in Experiment 1. Together with the items that remained the same, jokes did not statistically differ from the controls in comprehensibility ($b = 0.14$, $SE = 0.07$, $t = 1.85$, $p = .07$), but the jokes were still rated as funnier ($b = 1.07$, $SE = 0.09$, $t = 11.46$, $p < .001$), and had less predictable endings ($b = 0.66$, $SE = 0.08$, $t = 7.83$, $p < .001$) than the controls did. These changes to the material were preregistered.

4.6.1.3 Procedure.

Due to constraints in time and personnel, we reduced the number of filler trials from 480 to 336 (randomly dropped). As a result, two consecutive target trials were separated by five to nine fillers, with an average distance of seven (previously 10). All other aspects of the procedure remained the same as in Experiment 1. The entire experiment now took about 90 min. These changes to the procedure were also preregistered.

4.6.1.4 Data analysis.

Unless otherwise stated, there was no deviation from what was specified in the preregistration or from what was used in Experiment 1.

4.6.2 Results.

4.6.2.1 Recall Performance.

Similar to Experiment 1, when on-task, participants significantly more likely recalled a joke's ending than a non-funny control's ending, $b = 0.71$, $SE = 0.19$, $z = 3.71$, $p < .001$. This

recall advantage was again reduced during unintentional mind-wandering, $b = 0.47$, $SE = 0.26$, $z = 1.79$, $p = 0.07$, and intentional mind-wandering, $b = 0.28$, $SE = 0.33$, $z = 0.86$, $p = 0.39$. For the fourth contrast, the overall difference between on-task and mind-wandering was significant, $b = 1.14$, $SE = 0.15$, $z = 7.53$, $p < .001$ (refer to the supplemental material for the marginal means).

4.6.2.2 Eye Movement Measures.

Participants had more regressions-out from punchlines than from the controls' endings when they were on-task, $b = 0.36$, $SE = 0.11$, $z = 3.28$, $p = .001$. This difference was reduced during both unintentional mind-wandering, $b = 0.32$, $SE = 0.16$, $z = 1.98$, $p = 0.05$, and intentional mind-wandering, $b = 0.15$, $SE = 0.22$, $z = 0.69$, $p = 0.49$. Different from Experiment 1, the overall difference between on-task and mind-wandering was not significant, $b = 0.09$, $SE = 0.09$, $z = 0.99$, $p = 0.32$.

Participants had longer regression-path duration from punchlines than from neutral endings when they were on-task, $b = 0.13$, $SE = 0.04$, $t = 3.06$, $p = .003$. This difference was reduced during unintentional mind-wandering, $b = 0.11$, $SE = 0.06$, $t = 2.00$, $p = 0.05$. Interestingly, we found a somewhat larger estimate of the difference during intentional mind-wandering, $b = 0.16$, $SE = 0.08$, although it was only marginally significant, $t = 2.00$, $p = 0.05$. Finally, no significant difference was found between on-task and mind-wandering, $b < .001$, $SE = 0.03$, $t = 0.02$, $p = 0.99$, different from Experiment 1.

Similar to Experiment 1, we did not find any significant difference in gaze duration, $ps > .10$. Different from Experiment 1, however, the difference in total looking time between jokes and non-funny controls in the on-task condition was *not* significant, $b = 0.02$, $SE = 0.02$, $t = 0.81$, $p = 0.42$. The difference was also not significant during either unintentional mind-wandering or intentional mind-wandering, $ps > .05$.

Finally, we did not find any significant difference in skipping, $ps > .10$. In particular, the overall difference between on-task and mind-wandering was not significant, $b = -0.09$, $SE = 0.13$, $z = -0.70$, $p = 0.48$.

4.6.3 Discussion.

Despite some changes in stimuli and procedure, we observed significantly more rereading and better recall for joke endings compared to neutral endings when participants were on-task. These differences were generally reduced during both unintentional and intentional mind-wandering. Quite interestingly, there seemed to be a larger effect in regression–path duration between jokes and controls during intentional mind-wandering, compared to that when participants were on-task. Unlike regressions-out, duration measures treated skipping as a missing value instead of a zero. Thus, this difference only referred to cases in which the last word was fixated. Nevertheless, these results raised the possibility that, during intentional mind-wandering, the incongruity-resolution process was not always affected.

Different from Experiment 1, we did not observe a significant difference in total looking time when participants were on-task. In self-paced reading, where rereading is not permitted, reading time for punchlines is usually longer than that for neutral endings (Coulson & Kutas, 1998; Mayerhofer & Schacht, 2015). However, in free reading, the reader might not need to examine the punchline multiple times, as long as they had reread previous texts. The difference in total looking time was only marginally significant in another eye-tracking study that used a free reading paradigm (Coulson et al., 2006).

Results from the fourth contrast (non-mind-wandering vs. mind-wandering across all sentence types) differ from those in Experiment 1. We did not observe any significant difference in eye movement measures between mind-wandering and non-mind wandering, aggregating over

jokes and controls. Therefore, sentence endings, in general, received about the same amount of visual attention during mind-wandering and non-mind-wandering. Despite these inconsistencies, we again observed reduced rereading and recall advantage for jokes during mind-wandering.

4.7 Additional Analysis: Mind-Wandering and Lexical Processing

Existing theories offer different accounts for why deficits in higher-level linguistic processes occur during mind-wandering. The cascade model of inattention posits that deficits in higher-level processes are rooted in deficits in lower-level processes (Smallwood, 2011), whereas the levels of inattention hypothesis posits that higher-level deficits can still occur even when lower-level processes are intact (Schad et al., 2012). To adjudicate between the two accounts, we explored whether lexical processing at the punchline was also affected during mind-wandering. Specifically, we examined if the word frequency effect, as measured by the two early measures (gaze duration and skipping), was modulated by attention. If lexical processing at the ending was indeed impaired during mind-wandering, we should observe a smaller word frequency effect, compared to when participants were on-task. We combined data from Experiment 1 and the replication study to improve statistical power. This analysis was not preregistered.

The fixed effects of our analysis are shown in Figure 4-3. In general, the word frequency effect during mind-wandering did not significantly differ from that when participants were on-task, except for a smaller word frequency effect during intentional mind-wandering on word skipping. Thus, we did not find consistent evidence suggesting deficits at the lexical level during mind-wandering.

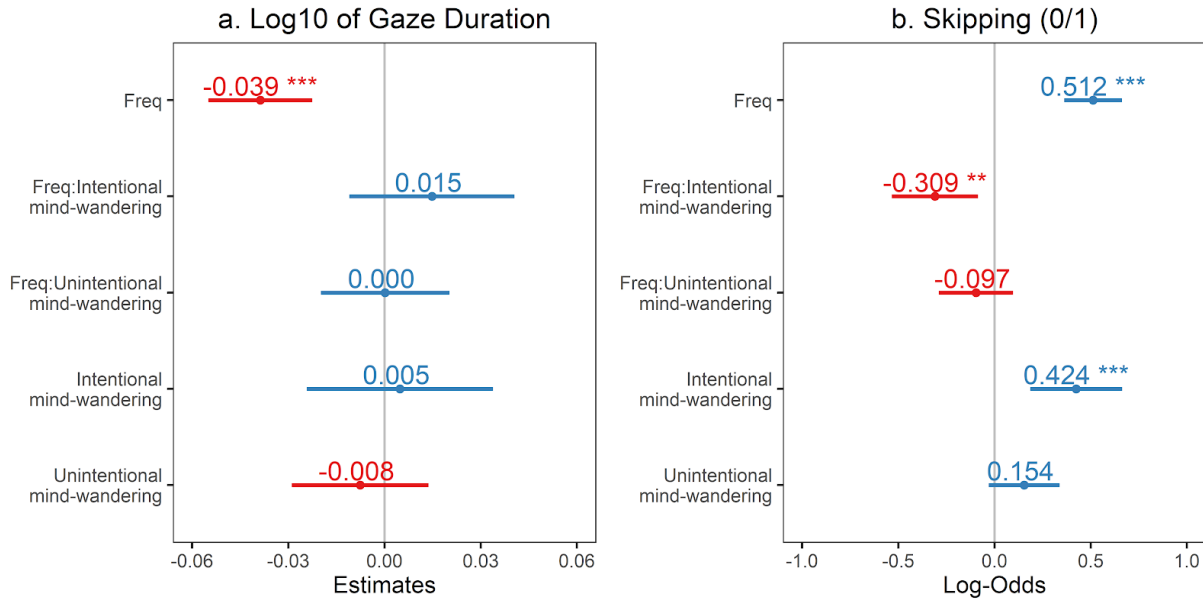


Figure 4-3. Fixed effects of regression analysis on the interaction between attention and word frequency. Attention (on-task, intentional mind-wandering, unintentional mind-wandering) was dummy-coded, with “on-task” as the reference level. Freq: Log10 of Word Frequency. *** $p < .001$, ** $p < .01$, * $p < .05$. Error bars show the 95% confidence intervals.

4.8 General Discussion

Garden-path jokes work by disrupting a narrative understanding built from the initial set-up. The higher-level processes of resolving the semantic incongruity are cognitively demanding, and this makes such jokes a promising venue for studying how we manage our attention in the face of distractions. The two studies described in the current article suggest that the resolution of semantic incongruity depends on the reader’s moment-to-moment attentional state.

4.8.1 The incongruity-resolution process.

Our results obtained from the on-task condition support the incongruity-resolution theory of garden-path joke processing (Suls, 1972, 1983). In both studies, jokes read without mind-wandering elicited more rereading from the punchline than from the non-funny controls, as if participants were reexamining the previous part of the text to find clues for an alternative explanation. Moreover, similar to Coulson et al. (2006), we did not observe any difference between jokes and controls in the early measures of reading (i.e., gaze duration and skipping),

but observed a significant difference in total looking time (only in Experiment 1). These findings suggest that the processing cost was related to a higher-level stage of language processing.

Additional rereading triggered by the punchline also fits with a recently updated computational model of eye movement control during reading (E-Z Reader 10; Reichle, Warren, & McConnell, 2009). In the E-Z reader model, the majority of regressions are due to difficulties in the post-lexical processing stage. Specifically, regressive eye movements can be initiated when the reader detects a failure in the integration of the current word into the overall meaning of the sentence (i.e., rapid integration failure). On the other hand, these results do not seem to fit well with the saccade generation with inhibition by foveal targets (SWIFT) model (Engbert, Nuthmann, Richter, & Kliegl, 2005). The SWIFT model assumes that the majority of regressions are due to unfinished lexical processing. Because garden-path jokes, in theory, do not entail additional processing at the lexical level, lexical difficulties do not seem to be the main reason that triggered rereading from punchlines. However, garden-path jokes can be a special case and our findings may have no bearing on the overall utility of the SWIFT model.

4.8.2 Joke processing during mind-wandering.

The current research contributes to a growing body of literature on how mind-wandering disrupts higher-level cognitive processes of reading. While some previous studies have used self-paced reading to answer this question (Smallwood, 2011; Smallwood et al., 2008), we reasoned that the rereading pattern in a free reading setting can convey important information about the reader's attentional state. Our two studies show that the additional rereading from the punchline observed in the on-task condition was generally reduced during mind-wandering. Mind-wandering also affected how well participants remembered the punchline during a subsequent cued-recall task. These results indicate that the incongruity-resolution process was impaired

during mind-wandering, making the processing of a joke less distinguishable from the processing of a neutral sentence.

Mind-wandering during reading has been theorized as “attentional decoupling,” such that attention shifts away from the linguistic input to internal thoughts and exerts less control on eye movements. Previous studies have shown that attentional decoupling can be measured at the lexical level, using variables such as word frequency (Foulsham et al., 2013; Reichle et al., 2010). The current study shows that attentional decoupling can also be measured at an advanced level of text processing. Moreover, our preliminary analysis did not find consistent evidence of deficits at the lexical level during mind-wandering. Thus, deficits at the higher-level stage during mind-wandering in our study cannot be solely attributed to deficits at the lexical level. This finding complicates assumptions that attentional decoupling during reading follows an “all-or-none” manner (Smallwood, 2011), but instead points to an alternative claim that attentional decoupling is graded in nature (Schad et al., 2012). Word recognition for skilled readers is largely automated, which may make it less susceptible to the effects of mind-wandering. However, higher-level processes are usually more effortful and may go astray during even weak levels of inattention.

We used cued-recall performance as an offline measure of joke processing. The results in the on-task condition replicated the humor effect, such that people have a better memory for information perceived as humorous (Schmidt, 1994, 2002). Importantly, our results suggest that one contributing factor is the elaborated visual processing triggered by semantic incongruity. However, this recall advantage disappeared during mind-wandering. While recall and comprehension are usually related, the current study did not directly measure whether the reader “got” the joke. Instead, we measured a cognitive process that is necessary but not sufficient for

getting a joke (Dyner, 2009). In other words, the reader might not have understood the joke after extensive processing. If so, the reader might still be able to report the ending but not the intended meaning of the text. In this sense, not getting a joke does not always mean that the reader was mind-wandering.

A potential future research direction is to look at whether rereading patterns, at least in some situations, can help detect mind-wandering. Research on mind-wandering has relied critically on asking participants to diagnose their mental states. While self-categorized mind-wandering seems valid (Smallwood & Schooler, 2015), the field requires more objective measurements to resolve important theoretical debates (Smallwood, 2013).

Moreover, the ability to identify when people are mind-wandering without interrupting them would open the possibility of systems that could respond to wandering attention in order to promote better task performance. There has been important progress in this line of research (e.g., Bixler & D’Mello, 2016; Faber et al., 2018). However, the best performing models appear to favor global features (text-irrelevant features) over local features. We note that local features may boost prediction performance in a more clearly defined setting, such as reading texts that contain occasional inconsistencies. When certain words trigger rereading, failing to do so can indicate a breakdown of attention.

We believe that mind-wandering research can benefit from connecting theories about attention to theories about language processing. To illustrate their interactions, eye-tracking will be an important methodology. We hope the current research will promote this integration so that we can better understand how people manage their attention in different contexts with different distractions that surround them.

Chapter 5 General Discussion

5.1 Summary of findings.

The current dissertation presented three studies that examined how MW affects visual attention in three distinct task settings: visual search, scene encoding, and reading comprehension. All three studies showed that MW was associated with decreased task performance at the behavioral level (e.g., response time, recognition, and recall). Importantly, eye-tracking provides a diagnostic tool to locate exactly the source of the performance cost. Specifically, study 1 (visual search) showed that MW, compared to being on-task, was associated with prolonged looking time on the search target if a salient distractor was also present. This extra processing time could explain the increased response time for distractor-present trials. Study 2 (scene encoding) showed that MW was associated with changes in the sequential aspect of eye movements, such that scanpaths were more repetitive and covered less area of the image during MW. Furthermore, fixation allocation during intentional MW, compared to being on-task, was less associated with what was meaningful in the scene. These eye movement changes could explain the worse recognition rate of stimuli presented during MW. Study 3 (reading) showed that MW was associated with reduced rereading from the punchline of garden-path jokes. This reduced rereading for jokes could explain the disappearance of the recall advantage of jokes compared to neutral sentences. Together, the results showed that MW is pervasive and detrimental in various task settings. Eye movements can provide novel insight into the underlying mechanisms of attention and eye movement control during MW.

5.2 Attentional decoupling during MW.

The results presented here also provide some clues regarding the nature of “attentional decoupling” during MW. It has been suggested that attentional decoupling entails a global reduction in processing visual input (Schooler et al., 2011) and those deficits in perceptual processing cause more serious deficits in higher-level processing (Smallwood, 2011). The current dissertation further qualifies this idea by illustrating that not all cognitive processes are similarly affected by attentional decoupling. First, in the visual search task, participants were able to resist the salient distractor even during MW, suggesting that they were somehow still sensitive to the perceptual information (at least the color information). Second, in the scene perception task, there was very limited evidence that the meaning-based guidance of attention was reduced during unintentional MW, which was the dominant form of MW in that task. Third, in the joke-reading task, the reduced re-reading of jokes occurred even though the word frequency effect was largely intact during MW. A common thread underlying these results is that cognitive processes that are largely effortless and automated might not be seriously affected (if at all) during MW. In the visual search task, the distractor suppression effect was well-practiced due to selection history; for scene perception, studies have shown that scene meaning guides visual attention automatically (Henderson & Hayes, 2017; Peacock et al., 2019); for reading, there is a lot of evidence showing that lexical processing (word frequency in particular) is highly automated and occurs rapidly for skilled readers (Pollatsek et al., 2006; Reingold, Reichle, Glaholt, & Sheridan, 2012; Sheridan & Reichle, 2016; White, Warrington, McGowan, & Paterson, 2015). The tasks we complete in daily life are usually supported by a range of cognitive processes that differ in their level of automaticity. Highly automated processes, such as the processing of scene semantics and word meaning, might be less vulnerable to the effects of

MW compared to less-practiced processes, such as detecting incongruity in sentences. By separating processes that are affected from those that are not affected, we can be more specific about creating interventions to reduce the effect of MW.

We should also keep in mind that MW is a heterogeneous concept. MW comes with many different forms and the different forms of MW might be associated with different underlying mechanisms. In this dissertation, there was evidence showing disruptions to automated processes (e.g., meaning-based guidance, word frequency effect) during intentional MW, but not unintentional MW. Indeed, if participants deliberately disengage from the task, then any task-related processing can be seriously affected, even for those that are effortless and automatic. By considering the heterogeneity of task-related and task-unrelated processing, we may develop a more nuanced view of task performance under distraction.

5.3 Task-general gaze measures of MW?

The three studies presented in this dissertation critically relied on local eye movement measures (e.g., looking time on target, re-reading, etc.) to investigate the effects of MW. However, local measures are task-dependent, and some (e.g., Faber et al., 2018) have argued that the objective detection of MW should be based on global measures (measures that can be computed independently from task settings). This section thus explores if there are generalizable global measures of MW across the three studies. Because intentional and unintentional MW were measured consistently throughout the studies, the potential differences between intentional and unintentional MW are also explored. The measures examined and their definitions are listed in Table 5-1.

Table 5-1. Global Eye Movement Measures Examined and Their Definitions.

Global Measures	Definitions
Fixation Count	The total number of fixations.
Fixation Duration Mean	The average duration of fixations.
Mu (μ)	Distributional parameter: mean component
Sigma (σ)	Distributional parameter: standard deviation component
Tau (τ)	Distributional parameter: exponential component
Fixation Duration Sum	The duration sum of fixations.
Fixation Dispersion	The root mean square of the Euclidean distance from each fixation to the average position of all fixations. It is reported on a 0-1 scale by normalizing on the maximum dispersion possible, with higher values indicating greater dispersion.
Area Covered	The total area covered by fixations (a circular region with a radius of 60 pixels) in proportion to the screen size.
Blink Count	The number of blinks.
Tracking Loss	The percentage of gaze signal not recorded by the eye-tracker.

It is worth noting that the distributional parameters were analyzed in addition to the overall mean of fixations. In Study 2 (visual search), the mean of fixation durations was larger during MW compared to being on-task. If increased fixation duration *is* a task-independent signature of MW, it is important to know if the underlying distributions are affected similarly across task settings. Like reaction times, the distribution of fixation durations typically follows the ex-Gaussian distribution (Staub, 2011; Staub, White, Drieghe, Hollway, & Rayner, 2010), which is a convolution of the normal distribution (defined by μ , the mean, and σ , the standard

deviation) and the exponential distribution (defined by τ , the exponential parameter). An increase in either μ or τ can elevate the overall mean, but the underlying mechanism is very different: an increase in μ indicates a fundamental change in the distribution where most fixations are shifted rightward, whereas an increase in τ indicates occasional disruptions in the distribution where only a small portion of fixations becomes more extreme with only little impact on the standard deviation σ . Separating the overall effect into μ and τ can reflect different kinds of effects that MW could have on visual processing. It has been shown that, while some factors can change both μ and τ (e.g., word frequency; Staub et al., 2010), some only affect μ (e.g., lexical predictability; Staub, 2011) or τ (e.g., music distraction, Zhang, Miller, Cleveland, & Cortina, 2018). Simply looking at the aggregated measures might conceal these effects. Thus, fixation durations were fitted to the ex-Gaussian distribution using the *retimes* package (Massidda, 2013) to see whether a change at the mean level corresponds to a change of μ or/and τ .

Data were analyzed by mixed-model ANOVAs using *lme4* (Bates et al., 2014). Study (visual search/scene encoding/reading [combined]) entered as a between-subject factor and attentional states (on-task/intentional MW/unintentional MW) entered as a within-subject factor. Significant interactions were further analyzed by pairwise comparisons of attentional states within each study (with Holm-Bonferroni correction).

5.3.1 Fixation count.

A mixed-effects ANOVA on fixation count showed a significant main effect of attentional states, $F(2, 306.54) = 7.27, p < .001$, which was qualified by a significant interaction, $F(4, 306.45) = 2.92, p = .02$. Post-hoc analysis results are shown in Table 5-2 and visualized in Figure 5-1. Compared to on-task episodes, MW was generally associated with fewer fixations

during reading and scene viewing, but not during visual search. Besides, intentional MW was associated with even fewer fixations compared to unintentional MW during reading comprehension.

Table 5-2. Pairwise comparisons on fixation count between attentional states within each study.

Comparisons		Estimate	<i>t</i> -scores	<i>p</i> -values
Study: Reading	On-task - Unintentional MW	0.8313	1.395	0.1641
	On-task - Intentional MW	2.6182	4.187	0.0001
	Unintentional MW - Intentional MW	1.7869	2.846	0.0095
Study: Scene Viewing	On-task - Unintentional MW	3.0742	4.040	0.0002
	On-task - Intentional MW	2.6774	2.860	0.0090
	Unintentional MW - Intentional MW	-0.3968	-0.421	0.6738
Study: Visual Search	On-task - Unintentional MW	-0.0239	-0.025	1.0000
	On-task - Intentional MW	0.1584	0.151	1.0000
	Unintentional MW - Intentional MW	0.1823	0.173	1.0000

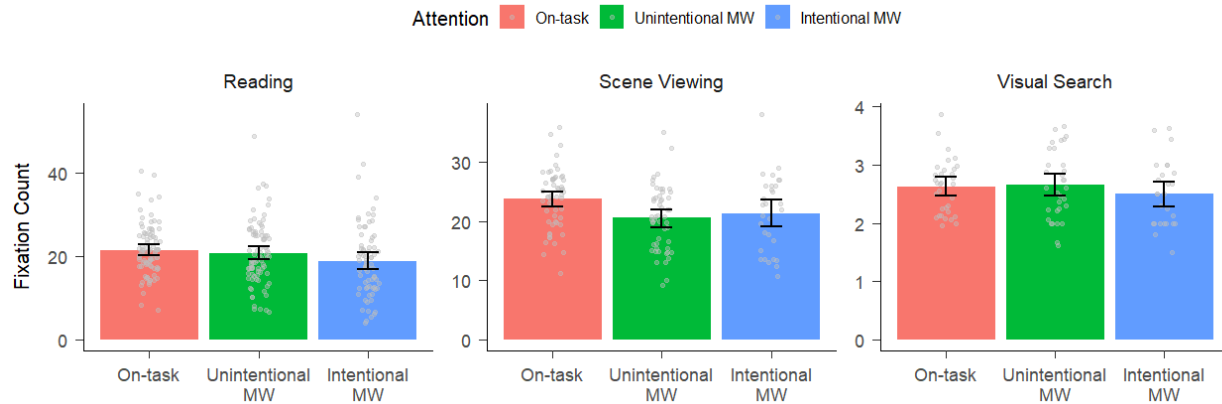


Figure 5-1. The number of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.3.2 Fixation duration.

A mixed-effects ANOVA on the mean of fixation durations showed a significant main effect of attentional states, $F(2, 301.53) = 17.25, p < .001$, which was qualified by a significant interaction, $F(4, 301.42) = 2.93, p = .02$. Post-hoc analysis results are shown in Table 5-3 and visualized in Figure 5-2. In general, MW was associated with longer fixations compared to on-task episodes. However, the differences reached significance only for unintentional MW during scene viewing, and intentional MW for scene viewing and visual search. Furthermore, intentional MW during the visual search task was associated with even longer fixations compared to unintentional MW.

Table 5-3. Pairwise comparisons on fixation duration between attentional states within each study.

	Comparisons	Estimate	<i>t</i> -scores	<i>p</i> -values
Study: Reading	On-task - Unintentional MW	-7.49	-1.280	0.4032

	On-task - Intentional MW	-10.25	-1.670	0.2877
	Unintentional MW - Intentional MW	-2.76	-0.448	0.6548
	On-task - Unintentional MW	-20.51	-2.748	0.0087
Study: Scene Viewing	On-task - Intentional MW	-46.99	-5.135	<.0001
	Unintentional MW - Intentional MW	-26.47	-2.873	0.0087
	On-task - Unintentional MW	-17.43	-1.819	0.1398
Study: Visual Search	On-task - Intentional MW	-30.27	-2.938	0.0107
	Unintentional MW - Intentional MW	-12.84	-1.242	0.2152

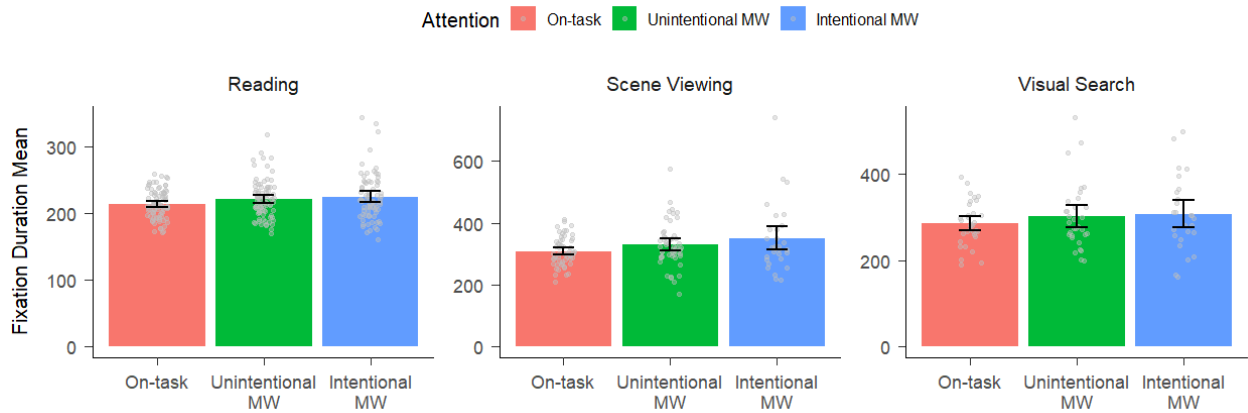


Figure 5-2. The mean duration of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.3.2.1 *Mu* (μ).

There was no main effect of attention $F(2, 303.23) = 1.68, p = .19$, or significant interaction between task and attention, $F(4, 302.29) = 2.23, p = .07$. Therefore, no post-hoc analysis was performed. See Figure 5-3 for a visualization of the data.

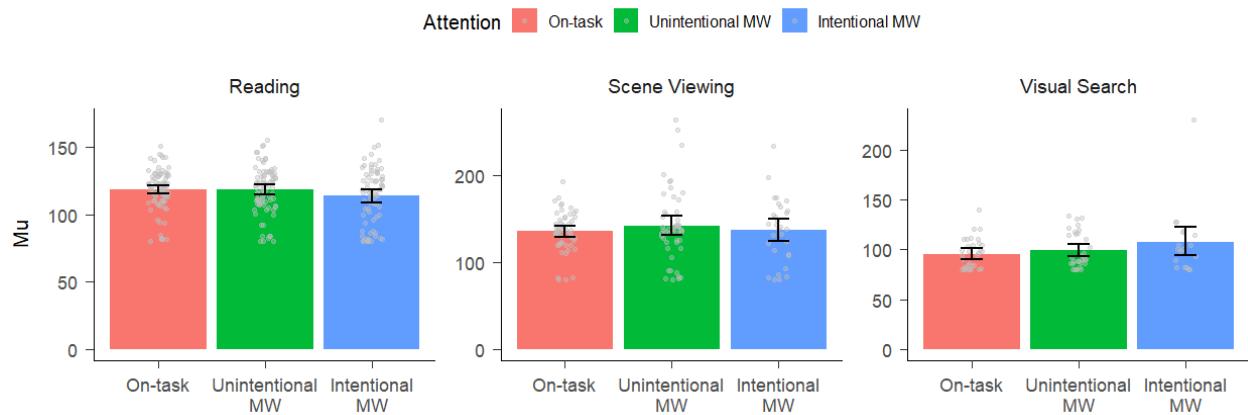


Figure 5-3. The estimated μ parameter of the distribution of fixation durations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.3.2.2 *Sigma* (σ).

There was a significant main effect of attention, $F(2, 327.30) = 24.31, p < .001$, which was qualified by a significant interaction between task and attention, $F(4, 325.69) = 23.95, p < .001$. Post-hoc analysis showed that the significant interaction came from differences between attentional states in the visual search task. Specifically, the sigma parameter was significantly larger during intentional MW compared to on-task, $b = 74.36, t = 9.51, p < .001$, and compared to unintentional MW, $b = 71.73, t = 9.09, p < .01$. All other comparisons were not significant, $ps > .69$.

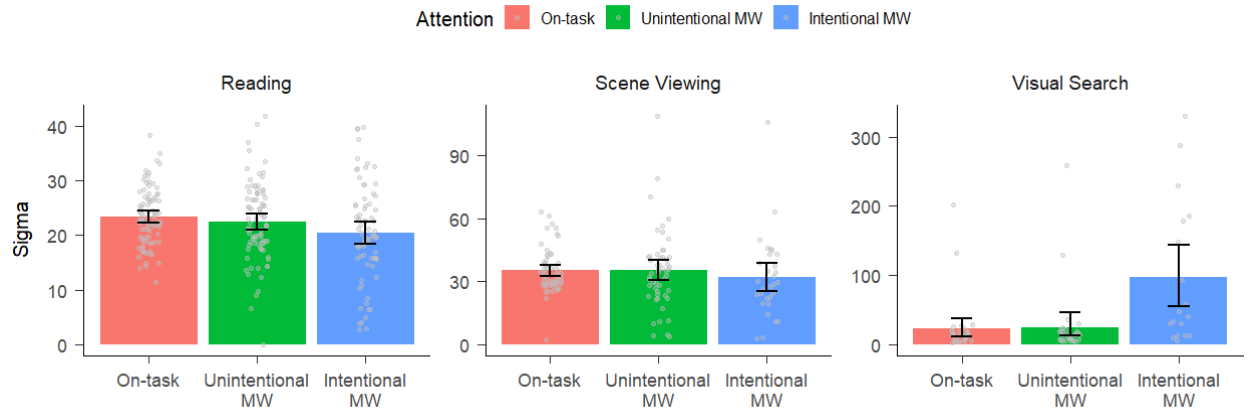


Figure 5-4. The estimated Sigma (σ) parameter of the distribution of fixation durations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.3.2.3 Tau (τ).

We found a significant main effect of attention, $F(2, 295.15) = 8.91, p < .001$. The interaction between attention and task was not significant, $F(4, 294.56) = .99, p = .41$. We went ahead and conducted a pairwise comparison between attentional states within each task. The results are shown in Table 5-4 and visualized in Figure 5-5. In general, the tau parameter during MW was larger compared to when on-task, but the differences were only significant for the reading task and the scene viewing task. Furthermore, the tau parameter for intentional MW was even larger compared to that in unintentional MW, which is consistent with the mean-level results.

Table 5-4. Pairwise comparisons on the tau parameter between attentional states within each study.

	Comparisons	Estimate	<i>t</i> -scores	<i>p</i> -values
Study: Reading	On-task - Unintentional MW	-6.84	-1.674	0.1766

	On-task - Intentional MW	-14.20	-3.312	0.0031
	Unintentional MW - Intentional MW	-7.36	-1.710	0.1766
	On-task - Unintentional MW	-9.19	-1.762	0.0791
Study: Scene Viewing	On-task - Intentional MW	-25.40	-3.959	0.0003
	Unintentional MW - Intentional MW	-16.21	-2.512	0.0250
	On-task - Unintentional MW	-8.36	-1.234	0.6545
Study: Visual Search	On-task - Intentional MW	-6.94	-0.836	0.8078
	Unintentional MW - Intentional MW	1.42	0.170	0.8650

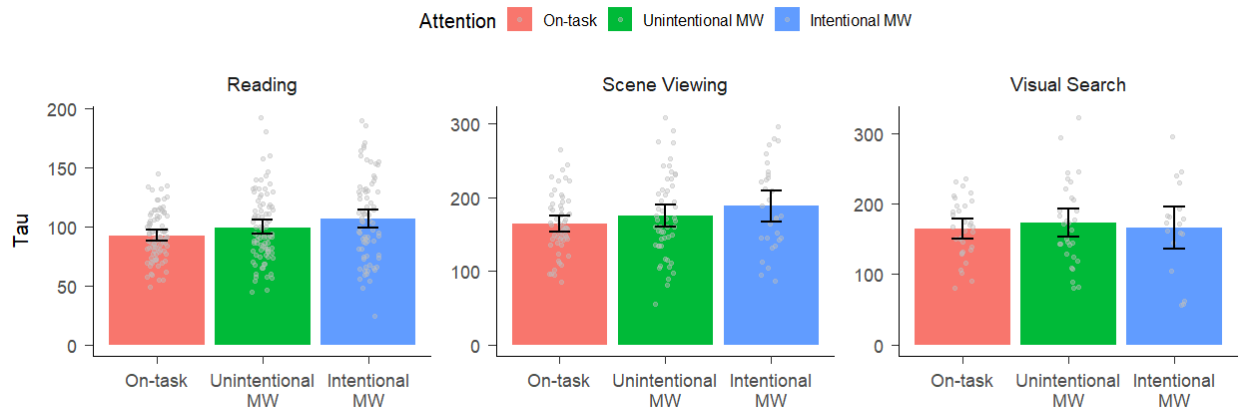


Figure 5-5. The estimated Tau (τ) parameter of the distribution of fixation durations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.3.3 Fixation duration sum.

The main effect of attention was not significant, $F(2, 309.52) = 2.56, p = .08$, but there was a significant attention-by-study interaction, $F(4, 309.43) = 3.17, p = .01$. Post-hoc analysis results are shown in Table 5-5 and visualized in Figure 5-6. In the reading task, intentional MW was associated with less total fixating time compared to on-task episodes; however, in the scene perception task, it was unintentional MW that was associated with less fixating time compared to on-task episodes.

Table 5-5. Pairwise comparisons on the sum of fixation durations between attentional states within each study.

	Comparisons	Estimate	<i>t</i> -scores	<i>p</i> -values
Study: Reading	On-task - Unintentional MW	35.9	0.259	0.7959
	On-task - Intentional MW	382.3	2.626	0.0273
	Unintentional MW - Intentional MW	346.4	2.369	0.0369
Study: Scene Viewing	On-task - Unintentional MW	646.3	3.648	0.0009
	On-task - Intentional MW	311.3	1.429	0.2547
	Unintentional MW - Intentional MW	-335.0	-1.529	0.2547
Study: Visual Search	On-task - Unintentional MW	-30.1	-0.133	1.0000

On-task - Intentional MW	-55.0	-0.225	1.0000
Unintentional MW - Intentional MW	-24.9	-0.102	1.0000

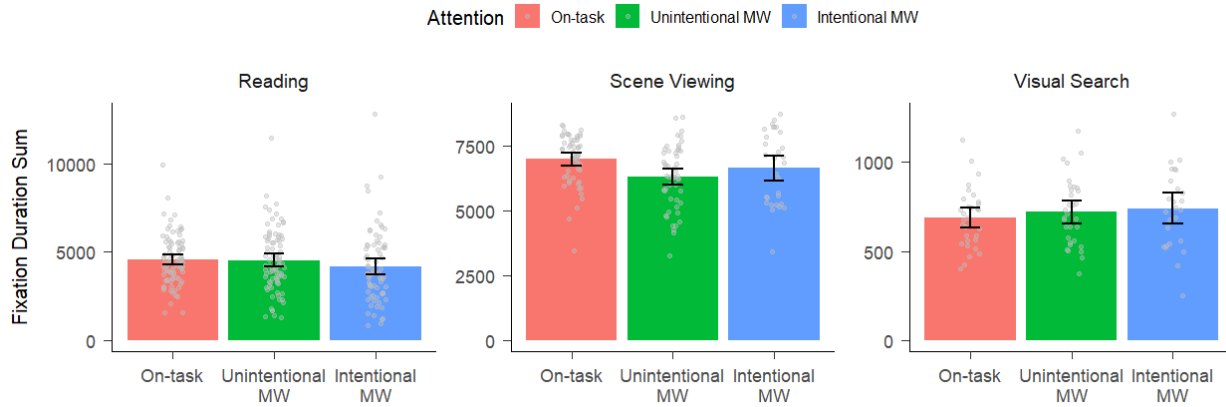


Figure 5-6. The sum duration of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.4.4 Fixation dispersion.

The main effect of attention was not significant, $F(2, 306.12) = 2.00, p = .14$, but there was a significant attention-by-study interaction, $F(4, 306.05) = 5.76, p < .001$. Post-hoc analysis results are shown in Table 5-6 and visualized in Figure 5-7. These results indicate that the significant interaction was driven by the reduced dispersion during MW in the scene perception task.

Table 5-6. Pairwise comparisons on the dispersion of fixations between attentional states within each study.

Comparisons		Estimate	<i>t</i> -scores	<i>p</i> -values
Study: Reading	On-task - Unintentional MW	-0.00321	-0.596	0.5517

	On-task - Intentional MW	-0.01198	-2.120	0.1044
	Unintentional MW - Intentional MW	-0.00878	-1.547	0.2460
	On-task - Unintentional MW	0.02071	3.010	0.0057
Study: Scene Viewing	On-task - Intentional MW	0.02961	3.490	0.0017
	Unintentional MW - Intentional MW	0.00889	1.043	0.2978
	On-task - Unintentional MW	-0.00853	-0.968	0.5692
Study: Visual Search	On-task - Intentional MW	0.01019	1.072	0.5692
	Unintentional MW - Intentional MW	0.01871	1.965	0.1510

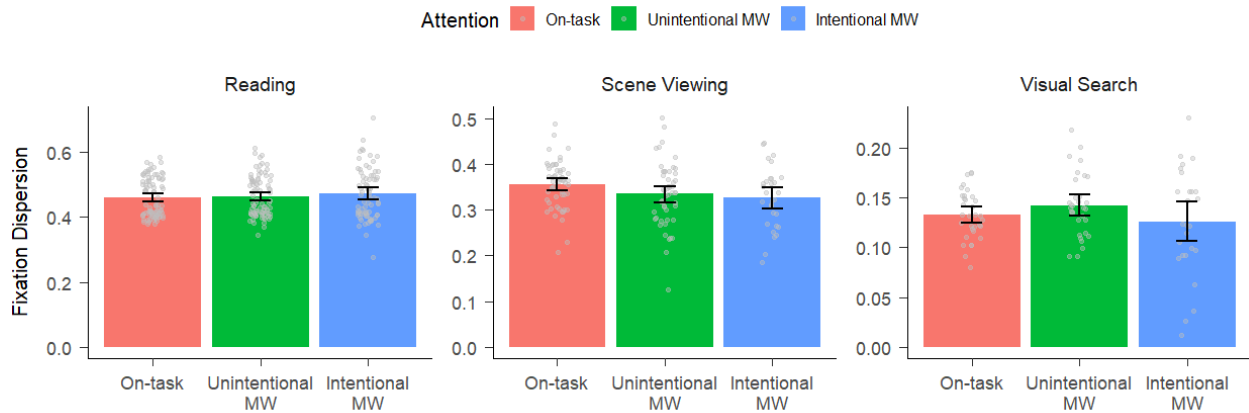


Figure 5-7. The dispersion of fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.4.5 Area covered.

The main effect of attention was significant, $F(2, 304.99) = 8.46, p < .001$, and there was a significant attention-by-study interaction, $F(4, 304.89) = 4.65, p = .001$. Post-hoc analysis results are shown in Table 5-7 and visualized in Figure 5-8. In general, there was less area coverage during MW in the reading task and the scene perception task, but the pattern was minimal in the visual search task. In the reading task, in particular, there was even less coverage of attention during intentional MW than during unintentional MW. These results are quite consistent with the fixation count results discussed in section 5.3.1.

Table 5-7. Pairwise comparisons on the percentage of area covered by fixations between attentional states within each study.

	Comparisons	Estimate	<i>t</i> -scores	<i>p</i> -values
Study: Reading	On-task - Unintentional MW	0.002793	0.661	0.5094
	On-task - Intentional MW	0.015477	3.490	0.0017
	Unintentional MW - Intentional MW	0.012684	2.849	0.0094
Study: Scene Viewing	On-task - Unintentional MW	0.027929	5.177	<.0001
	On-task - Intentional MW	0.023613	3.563	0.0008
	Unintentional MW - Intentional MW	-0.004316	-0.647	0.5181

	On-task - Unintentional MW	0.000308	0.044	1.0000
Study: Visual Search	On-task - Intentional MW	0.001723	0.231	1.0000
	Unintentional MW - Intentional MW	0.001416	0.190	1.0000

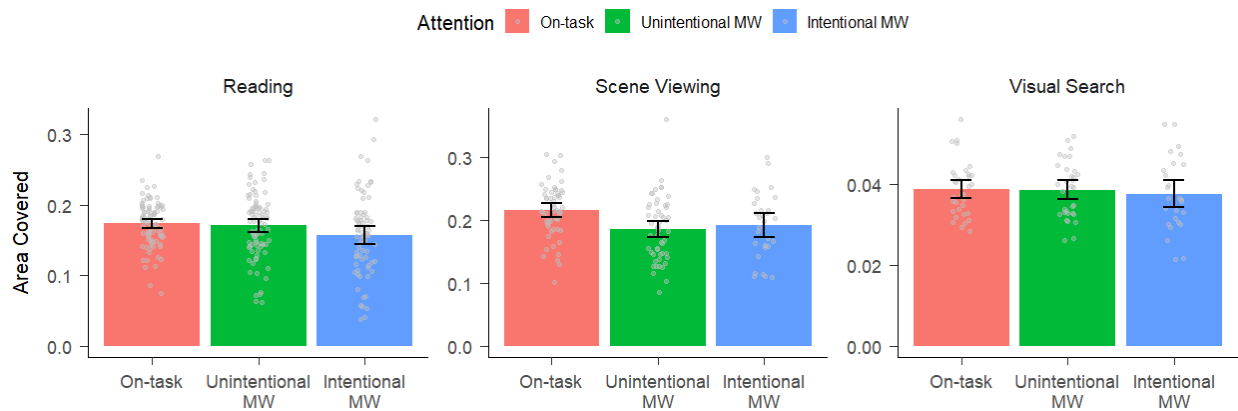


Figure 5-8. The percentage of area covered by fixations during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.4.6 Blink count.

The main effect of attention was significant, $F(2, 301.18) = 5.09, p = .007$, but the attention-by-study interaction was not significant, $F(4, 301.12) = .44, p = .78$. We went on to conduct a set of pairwise comparisons within each study. The only significant result comparison shows that there were more blinks during unintentional MW than during on-task episodes in the scene perception task, $b = .41, t = 3.01, p = .008$. All other $ps > .13$.

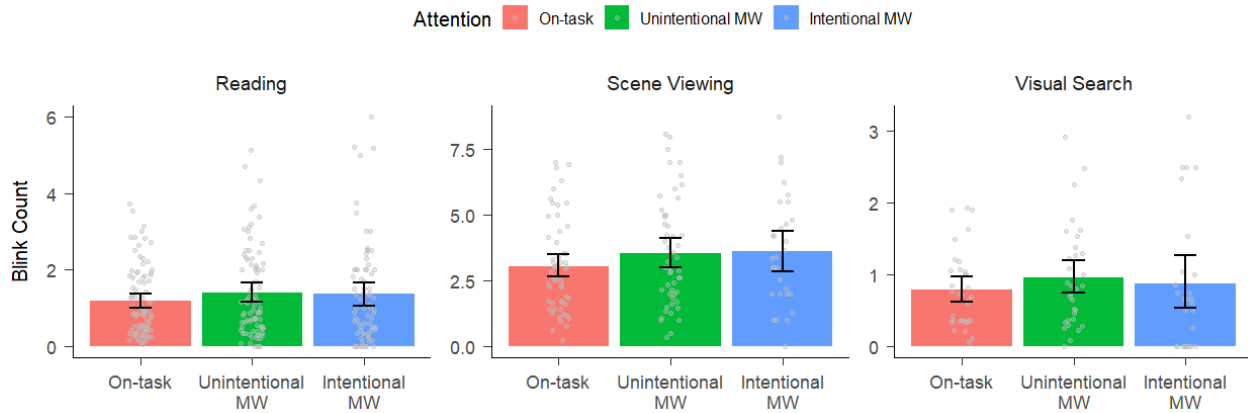


Figure 5-9. The number of blinks during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

5.4.7 Tracking loss.

The main effect of attention was significant, $F(2, 308.14) = 6.12, p = .002$, but the attention-by-study interaction was not significant, $F(4, 308.04) = 1.91, p = .11$. As done previously, we went on to conduct a set of pairwise comparisons within each study. The results are shown in Table 5-8 and visualized in Figure 5-10. In general, MW was associated with greater tracking loss, although this effect was relatively larger for the scene viewing task and specifically for unintentional MW.

Table 5-8. Pairwise comparisons on tracking loss between attentional states within each study.

Comparisons	Estimate	<i>t</i> -scores	<i>p</i> -values
On-task - Unintentional MW	-0.01422	-1.835	0.1351
Study: Reading			
On-task - Intentional MW	-0.01828	-2.248	0.0759

	Unintentional MW - Intentional MW	-0.00406	-0.497	0.6194
	On-task - Unintentional MW	-0.02608	-2.636	0.0176
Study: Scene Viewing	On-task - Intentional MW	0.01020	0.840	0.4018
	Unintentional MW - Intentional MW	0.03628	2.966	0.0097
	On-task - Unintentional MW	-0.02104	-1.658	0.2952
Study: Visual Search	On-task - Intentional MW	-0.01041	-0.762	0.8770
	Unintentional MW - Intentional MW	0.01062	0.776	0.8770

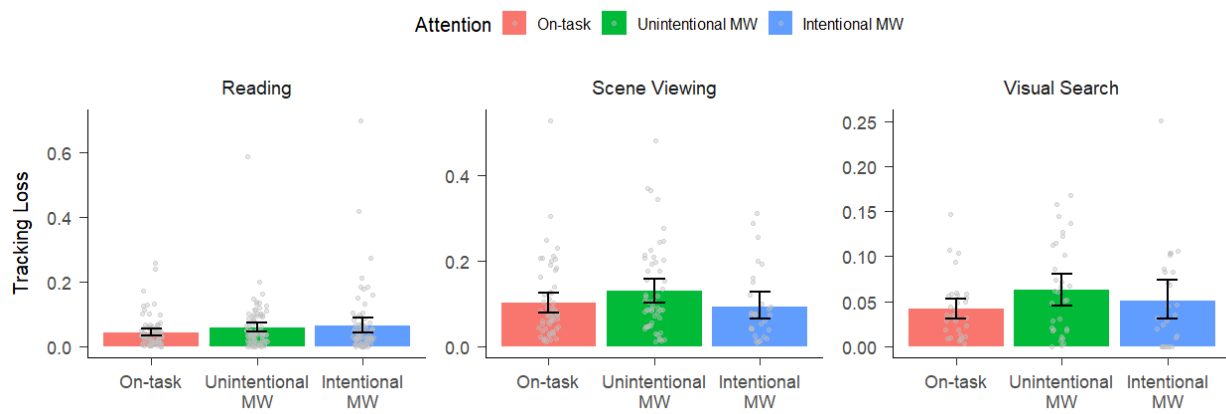


Figure 5-10. The percentage of tracking loss during on-task, intentional MW, and unintentional MW episodes in three task settings. Error bars showed 95% confidence intervals. Gray dots represent subject-level means.

As evident from these results, no single eye movement measure successfully generalized to the three tasks. However, based on these results we can still draw several tentative conclusions regarding eye movements of MW across tasks.

- (1) **MW is associated with a reduced amount of visual processing in tasks that require extensive sampling of the stimuli.** There were fewer fixations during MW in the reading task and the scene perception task but not the visual search task. The results of another two measures, the area covered by fixations and the sum of fixation durations, showed similar patterns. Reading and scene perception require the viewer to allocate a sufficient number of fixations to extract necessary information. However, in the visual search task, participants were asked to locate the target as quickly as possible while ignoring other non-targets. Too many fixations may suggest worse performance. Therefore, in this task setting, both being on-task and MW might be associated with a small number of fixations.
- (2) **MW is associated with an increase in extremely long fixations in reading and scene perception.** The distributional analyses on the reading and the scene viewing data showed that the right tail of the distribution of fixation durations was longer during MW than during on-task episodes. The results of the reading task are especially interesting in that there were significant changes in the *tau* parameter despite that there was no overall mean difference. These results suggest that there are more processing disruptions during MW, causing long delays in eye movements. In contrast, there were no significant differences in the *mu* parameter between on-task and MW states. An increase in the *mu* parameter is often seen in processing difficulty manipulations such as changes in word frequency and word

predictability (Staub, 2011; Staub et al., 2010). On the other hand, an increase in the τ parameter was found with an increased level of external distraction (Zhang et al., 2018). Here, the results showed that an increase in the τ parameter could also be associated with internal distraction. As for the visual search task, we did find a difference in the overall mean, but there was no difference in either μ or τ . It should be noted that the estimation of distributional parameters might be less accurate for this task because there were not so many fixations available.

- (3) **MW might be associated with worse tracking quality.** There appears to be a small increase in tracking loss during MW in reading and scene perception, but not in visual search. The loss of signal may be related to the reduction in the number of fixations parsed by the eye-tracker, as mentioned in (1). Many reasons can cause a loss of signal, such as head movement, the closing of eyelids, looking away from the screen, etc. – all of these may tend to occur during MW. Why did the reduced tracking quality not happen during visual search? It might be because of the 2-second fixation cross imposed at the beginning of each trial – participants must look at the cross for 2 seconds to trigger the search array, and this procedure might have improved tracking quality during MW.
- (4) **Intentional and unintentional MW showed similar trends.** In the reading task, intentional MW was associated with fewer fixations, less covered area, and a smaller fixation duration sum compared to unintentional MW. But this pattern did not hold for scene perception and visual search. The differences in eye movements between intentional and unintentional MW may be localized to

specific task scenarios and are not easily generalizable. However, it is worth pointing out that, in almost every case, intentional and unintentional MW did not differ in their trends with respect to the on-task state. For example, while there was no difference between intentional and unintentional MW on the number of fixations in the scene perception task, both were associated with fewer fixations when compared to on-task episodes.

5.4 Concluding remarks.

How can we use the study of distraction to improve human cognition? By looking at how distraction disrupts cognitive processes in different task scenarios, we can gain insight into the basic mechanisms of attention so we can come up with ideas to restore attention. Moreover, oftentimes people are not aware that they are being distracted. A better understanding of the science of MW provides a way around the fundamental paradox of MW research, namely the fact that we are asking people to pay attention to the fact that they are not paying attention. As studies uncover more objective signatures of MW, we may be able to use such measures to identify MW states in real time independent of subjective reports. However, it is critical to understand that MW may have different effects on eye movements depending on the task characteristics. In order to formulate a unified understanding of what MW is, we must first understand the diverse ways MW manifests itself in different tasks. Understanding how gaze parameters change as a function of attentional states is an important step in optimizing human attention. I hope the current dissertation will be useful in helping us achieve this ultimate goal, while respecting the complexity and diversity of human cognition.

Bibliography

- Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2015). A comparison of scanpath comparison methods. *Behavior Research Methods*, *47*(4), 1377–1392. <https://doi.org/10.3758/s13428-014-0550-3>
- Anderson, N. C., Bischof, W. F., Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, *45*(3), 842–856. <https://doi.org/10.3758/s13428-012-0299-5>
- Anderson, N. C., Ort, E., Kruijne, W., Meeter, M., & Donk, M. (2015). It depends on *when* you look at it: Saliency influences eye movements in natural scene viewing and search early in time. *Journal of Vision*, *15*(5), 9. <https://doi.org/10.1167/15.5.9>
- Antrobus, J. S. (1968). Information theory and stimulus-independent thought. *British Journal of Psychology*, *59*(4), 423–430.
- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*(1), 403–450. <https://doi.org/10.1146/annurev.neuro.28.061604.135709>
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. <https://doi.org/10.1016/j.tics.2012.06.010>
- Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*(5), 485–496. <https://doi.org/10.3758/BF03205306>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*. <https://doi.org/10.1016/j.jml.2012.11.001>
- 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>
- Bastian, M., & Sackur, J. (2013). Mind wandering at the fingertips: Automatic parsing of subjective states based on response time variability. *Frontiers in Psychology*, *4*. <https://doi.org/10.3389/fpsyg.2013.00573>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Benarroch, E. E. (2009). The locus ceruleus norepinephrine system: Functional organization and potential clinical significance. *Neurology*, *73*(20), 1699–1704. <https://doi.org/10.1212/WNL.0b013e3181c2937c>
- Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus–noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, *42*(1), 33–84. [https://doi.org/10.1016/S0165-0173\(03\)00143-7](https://doi.org/10.1016/S0165-0173(03)00143-7)
- Bicknell, K., & Levy, R. (2011). Why readers regress to previous words: A statistical analysis. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *7*.

- Bixler, R., & D’Mello, S. (2015). Automatic gaze-based detection of mind wandering with metacognitive awareness. *International Conference on User Modeling, Adaptation, and Personalization*, 31–43. Springer.
- Bixler, R., & D’Mello, S. (2016). Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Modeling and User-Adapted Interaction*, 26(1), 33–68. <https://doi.org/10.1007/s11257-015-9167-1>
- Bloom, B. S. (1953). Thought-processes in lectures and discussions. *The Journal of General Education*, 7(3), 160–169.
- Braboszcz, C., & Delorme, A. (2011). Lost in thoughts: Neural markers of low alertness during mind wandering. *NeuroImage*, 54(4), 3040–3047. <https://doi.org/10.1016/j.neuroimage.2010.10.008>
- Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, 21(1), 1–16.
- Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, 89(3–4), 319–325. <https://doi.org/10.1007/s00421-003-0807-5>
- Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, 111(1), 98–113. <https://doi.org/10.1016/j.cognition.2008.12.009>
- 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>
- Christoff, K., Mills, C., Andrews-Hanna, J. R., Irving, Z. C., Thompson, E., Fox, K. C. R., & Kam, J. W. Y. (2018). Mind-Wandering as a Scientific Concept: Cutting through the Definitional Haze. *Trends in Cognitive Sciences*, 22(11), 957–959. <https://doi.org/10.1016/j.tics.2018.07.004>
- Conners, C. K., Erhardt, D., & Sparrow, E. P. (1999). *Conners’ adult ADHD rating scales (CAARS): Technical manual*. MHS North Tonawanda.
- Coulson, S., & Kutas, M. (1998). Frame-shifting and sentential integration. *Cognitive Science Technical Report*, 98(02).
- Coulson, S., Urbach, T. P., & Kutas, M. (2006). Looking back: Joke comprehension and the space structuring model. *Humor*. <https://doi.org/10.1515/HUMOR.2006.013>
- Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods*, 42(3), 692–700. <https://doi.org/10.3758/BRM.42.3.692>
- Dalmajer, E. S., Mathôt, S., & der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, 46(4), 913–921.
- Davies, M. (2008). *The corpus of contemporary American English*. BYU, Brigham Young University.

- deBettencourt, M. T., Keene, P. A., Awh, E., & Vogel, E. K. (2019). Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour*, 3(8), 808–816. <https://doi.org/10.1038/s41562-019-0606-6>
- Dewhurst, R., Foulsham, T., Jarodzka, H., Johansson, R., Holmqvist, K., & Nyström, M. (2018). How task demands influence scanpath similarity in a sequential number-search task. *Vision Research*, 149, 9–23.
- Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods*, 44(4), 1079–1100.
- Droll, J. A., & Hayhoe, M. M. (2007). Trade-offs between gaze and working memory use. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 1352–1365. <https://doi.org/10.1037/0096-1523.33.6.1352>
- Dynel, M. (2009). *Humorous garden-paths: A pragmatic-cognitive study*. Cambridge Scholars Publishing.
- Ehrlichman, H., & Micic, D. (2012). Why Do People Move Their Eyes When They Think? *Current Directions in Psychological Science*, 21(2), 96–100. <https://doi.org/10.1177/0963721412436810>
- Ehrlichman, H., Micic, D., Sousa, A., & Zhu, J. (2007). Looking for answers: Eye movements in non-visual cognitive tasks. *Brain and Cognition*, 64(1), 7–20. <https://doi.org/10.1016/j.bandc.2006.10.001>
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review*, 112(4), 777–813. <https://doi.org/10.1037/0033-295X.112.4.777>
- Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the Zone or Zoning Out? Tracking Behavioral and Neural Fluctuations During Sustained Attention. *Cerebral Cortex*, 23(11), 2712–2723. <https://doi.org/10.1093/cercor/bhs261>
- Faber, M., Bixler, R., & D’Mello, S. K. (2018). An automated behavioral measure of mind wandering during computerized reading. *Behavior Research Methods*, 50(1), 134–150. <https://doi.org/10.3758/s13428-017-0857-y>
- Folk, C. L., & Remington, R. (1998). Selectivity in Distraction by Irrelevant Featural Singletons: Evidence for Two Forms of Attentional Capture. *Journal of Experimental Psychology: Learning Memory and Cognition*, 24(3), 847–858.
- Forster, S., & Lavie, N. (2014). Distracted by your mind? Individual differences in distractibility predict mind wandering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(1), 251–260. <https://doi.org/10.1037/a0034108>
- Forster, S., Robertson, D. J., Jennings, A., Asherson, P., & Lavie, N. (2013). Plugging the attention deficit: Perceptual load counters increased distraction in ADHD. *Neuropsychology*, 28(1), 91. <https://doi.org/10.1037/neu0000020>
- Foulsham, T., Farley, J., & Kingstone, A. (2013). Mind wandering in sentence reading: Decoupling the link between mind and eye. *Canadian Journal of Experimental Psychology*. <https://doi.org/10.1037/a0030217>
- Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, 8(2), 6. <https://doi.org/10.1167/8.2.6>
- Frank, D. J., Nara, B., Zavagnin, M., Touron, D. R., & Kane, M. J. (2015). Validating older adults’ reports of less mind-wandering: An examination of eye movements and

- dispositional influences. *Psychology and Aging*, 30(2), 266–278.
<https://doi.org/10.1037/pag0000031>
- Franklin, M. S. (2013). Window to the wandering mind: Pupillometry of spontaneous thought while reading. *THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY*, 7.
- Franklin, M. S., Mooneyham, B. W., Baird, B., & Schooler, J. W. (2014). Thinking one thing, saying another: The behavioral correlates of mind-wandering while reading aloud. *Psychonomic Bulletin & Review*, 21(1), 205–210. <https://doi.org/10.3758/s13423-013-0468-2>
- Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review*, 18(5), 992–997. <https://doi.org/10.3758/s13423-011-0109-6>
- Frazier, L. (1998). Getting There (Slowly). *Journal of Psycholinguistic Research*, 27(2), 123–146. <https://doi.org/10.1023/A:1023241830722>
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210. [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1)
- Gaspelin, N., Gaspar, J. M., & Luck, S. J. (2019). Oculomotor inhibition of salient distractors: Voluntary inhibition cannot override selection history. *Visual Cognition*, 27(3–4), 227–246. <https://doi.org/10.1080/13506285.2019.1600090>
- Gaspelin, N., Leonard, C. J., & Luck, S. J. (2015). Direct Evidence for Active Suppression of Salient-but-Irrelevant Sensory Inputs. *Psychological Science*, 26(11), 1740–1750. <https://doi.org/10.1177/0956797615597913>
- Gaspelin, N., Leonard, C. J., & Luck, S. J. (2017). Suppression of overt attentional capture by salient-but-irrelevant color singletons. *Attention, Perception, & Psychophysics*, 79(1), 45–62. <https://doi.org/10.3758/s13414-016-1209-1>
- Gaspelin, N., & Luck, S. J. (2018a). Distinguishing among potential mechanisms of singleton suppression. *Journal of Experimental Psychology: Human Perception and Performance*, 44(4), 626–644. <https://doi.org/10.1037/xhp0000484>
- Gaspelin, N., & Luck, S. J. (2018b). The Role of Inhibition in Avoiding Distraction by Salient Stimuli. *Trends in Cognitive Sciences*, 22(1), 79–92. <https://doi.org/10.1016/j.tics.2017.11.001>
- Giambra, L. M., & Grodsky, A. (1989). Task-Unrelated Images and Thoughts While Reading. In J. E. Shorr, P. Robin, J. A. Connella, & M. Wolpin (Eds.), *Imagery: Current Perspectives* (pp. 27–31). https://doi.org/10.1007/978-1-4899-0876-6_3
- Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, 10(19), 1209–1212. [https://doi.org/10.1016/S0960-9822\(00\)00729-6](https://doi.org/10.1016/S0960-9822(00)00729-6)
- Gilzenrat, M. S., Cohen, J. D., Rajkowski, J., & Aston-Jones, G. (2003). Pupil dynamics predict changes in task engagement mediated by locus coeruleus. *Society for Neuroscience Abstracts*, 515, 19.
- Gilzenrat, M. S., Nieuwenhuis, S., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, 10(2), 252–269. <https://doi.org/10.3758/CABN.10.2.252>
- Golchert, J., Smallwood, J., Jefferies, E., Seli, P., Huntentburg, J. M., Liem, F., ... Margulies, D. S. (2017). Individual variation in intentionality in the mind-wandering state is reflected in

- the integration of the default-mode, fronto-parietal, and limbic networks. *NeuroImage*, 146, 226–235. <https://doi.org/10.1016/j.neuroimage.2016.11.025>
- Grandchamp, R., Braboszcz, C., & Delorme, A. (2014). Oculometric variations during mind wandering. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00031>
- Gurtner, L. M., Bischof, W. F., & Mast, F. W. (2019). Recurrence quantification analysis of eye movements during mental imagery. *Journal of Vision*, 19(1), 17. <https://doi.org/10.1167/19.1.17>
- Häkkinen, H., LicPsych, Summala, H., Partinen, M., Tiihonen, M., & Silvo, J. (1999). Blink Duration as an Indicator of Driver Sleepiness in Professional Bus Drivers. *Sleep*, 22(6), 798–802. <https://doi.org/10.1093/sleep/22.6.798>
- Harel, J., Koch, C., & Perona, P. (2007). Graph-based visual saliency. *Advances in Neural Information Processing Systems*, 545–552.
- Hayes, T. R., & Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*, 48(2), 510–527.
- He, J., Becic, E., Lee, Y. C., & McCarley, J. S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human Factors*, 53(1), 13–21. <https://doi.org/10.1177/0018720810391530>
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7(11), 498–504.
- Henderson, J. M., Brockmole, J. R., Castelano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In *Eye movements* (pp. 537–III). Elsevier.
- Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour*, 1(10), 743–747. <https://doi.org/10.1038/s41562-017-0208-0>
- Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review*, 16(5), 850–856.
- Hershman, R., Henik, A., & Cohen, N. (2018). A novel blink detection method based on pupillometry noise. *Behavior Research Methods*, 50(1), 107–114. <https://doi.org/10.3758/s13428-017-1008-1>
- Hutt, S., Mills, C., Bosch, N., Krasich, K., Brockmole, J., & D’Mello, S. (2017). “Out of the Fr-Eye-ing Pan”: Towards Gaze-Based Models of Attention during Learning with Technology in the Classroom. *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization - UMAP '17*, 94–103. <https://doi.org/10.1145/3079628.3079669>
- Hutt, S., Mills, C., White, S., Donnelly, P. J., & D’Mello, S. K. (2016). The Eyes Have It: Gaze-Based Detection of Mind Wandering during Learning with an Intelligent Tutoring System. *International Educational Data Mining Society*.
- Ishihara. (2010). *Ishihara’s Tests for Colour Deficiency 2010: 14 Plate Book Concise Edition* (1st ed.). Retrieved from <https://www.amazon.com/Ishihara-Chart-Books-Color-Deficiency/dp/B000FLTPZU>
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10–12), 1489–1506.
- Jubera-García, E., Gevers, W., & Van Opstal, F. (2019). Influence of content and intensity of thought on behavioral and pupil changes during active mind-wandering, off-focus, and

- on-task states. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-019-01865-7>
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585.
- Kam, J. W., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W., & Handy, T. C. (2011). Slow fluctuations in attentional control of sensory cortex. *Journal of Cognitive Neuroscience*, *23*(2), 460–470.
- 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: An experience-sampling study of working memory and executive control in daily life. *Psychological Science*, *18*(7), 614–621.
- Kane, M. J., & McVay, J. C. (2012). What Mind Wandering Reveals About Executive-Control Abilities and Failures. *Current Directions in Psychological Science*, *21*(5), 348–354. <https://doi.org/10.1177/0963721412454875>
- Kane, M. J., Smeekens, B. A., von Bastian, C. C., Lurquin, J. H., Carruth, N. P., & Miyake, A. (2017). A combined experimental and individual-differences investigation into mind wandering during a video lecture. *Journal of Experimental Psychology: General*, *146*(11), 1649–1674. <https://doi.org/10.1037/xge0000362>
- Killingsworth, M. A., & Gilbert, D. T. (2010). A Wandering Mind Is an Unhappy Mind. *Science*, *330*(6006), 932–932. <https://doi.org/10.1126/science.1192439>
- Kinsbourne, M. (1972). Eye and Head Turning Indicates Cerebral Lateralization. *Science*, *176*(4034), 539–541. <https://doi.org/10.1126/science.176.4034.539>
- Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1–2), 262–284. <https://doi.org/10.1080/09541440340000213>
- Klinger, E., & Cox, W. M. (1987). Dimensions of Thought Flow in Everyday Life. *Imagination, Cognition and Personality*, *7*(2), 105–128. <https://doi.org/10.2190/7K24-G343-MTQW-115V>
- Konishi, M., Brown, K., Battaglini, L., & Smallwood, J. (2017). When attention wanders: Pupillometric signatures of fluctuations in external attention. *Cognition*, *168*, 16–26. <https://doi.org/10.1016/j.cognition.2017.06.006>
- Kopp, K., D’Mello, S., & Mills, C. (2015). Influencing the occurrence of mind wandering while reading. *Consciousness and Cognition*, *34*, 52–62. <https://doi.org/10.1016/j.concog.2015.03.003>
- Krasich, K., McManus, R., Hutt, S., Faber, M., D’Mello, S. K., & Brockmole, J. R. (2018). Gaze-Based Signatures of Mind Wandering During Real-World Scene Processing. *Journal of Experimental Psychology: General*, *147*(8), 1111–1124. <https://doi.org/10.1037/xge0000411>
- Kristjansson, S. D., Stern, J. A., Brown, T. B., & Rohrbaugh, J. W. (2009). Detecting phasic lapses in alertness using pupillometric measures. *Applied Ergonomics*, *40*(6), 978–986. <https://doi.org/10.1016/j.apergo.2009.04.007>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). <https://doi.org/10.18637/jss.v082.i13>
- Kvavilashvili, L., & Mandler, G. (2004). Out of one’s mind: A study of involuntary semantic memories. *Cognitive Psychology*, *48*(1), 47–94.

- Lakens, D., & Caldwell, A. R. (2019). *Simulation-Based Power-Analysis for Factorial ANOVA Designs*. <https://doi.org/10.31234/osf.io/baxsf>
- Lamy, D., & Tsal, Y. (1999). A salient distractor does not disrupt conjunction search. *Psychonomic Bulletin & Review*, *6*(1), 93–98. <https://doi.org/10.3758/BF03210814>
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25–26), 3559–3565.
- Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, *13*(1), 132–138. <https://doi.org/10.3758/BF03193824>
- Lenartowicz, A., Simpson, G. V., & Cohen, M. S. (2013). Perspective: Causes and functional significance of temporal variations in attention control. *Frontiers in Human Neuroscience*, *7*. <https://doi.org/10.3389/fnhum.2013.00381>
- Maffei, A., & Angrilli, A. (2018). Spontaneous eye blink rate: An index of dopaminergic component of sustained attention and fatigue. *International Journal of Psychophysiology*, *123*, 58–63.
- Mason, D. J., Humphreys, G. W., & Kent, L. (2005). Insights into the control of attentional set in ADHD using the attentional blink paradigm. *Journal of Child Psychology and Psychiatry*, *46*(12), 1345–1353. <https://doi.org/10.1111/j.1469-7610.2005.01428.x>
- Massidda, D. (2013). *retimes: Reaction Time Analysis*. Retrieved from <https://cran.r-project.org/package=retimes>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.
- Mayerhofer, B., & Schacht, A. (2015). From incoherence to mirth: Neuro-cognitive processing of garden-path jokes. *Frontiers in Psychology*, *6*(MAY). <https://doi.org/10.3389/fpsyg.2015.00550>
- McIntyre, N. A., & Foulsham, T. (2018). Scanpath analysis of expertise and culture in teacher gaze in real-world classrooms. *Instructional Science*. <https://doi.org/10.1007/s11251-017-9445-x>
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 196–204. <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>
- Meghanathan, R. N., Nikolaev, A. R., & van Leeuwen, C. (2019). Refixation patterns reveal memory-encoding strategies in free viewing. *Attention, Perception, & Psychophysics*, *81*(7), 2499–2516. <https://doi.org/10.3758/s13414-019-01735-2>
- Mills, C., Bixler, R., Wang, X., & D’Mello, S. (2015). Automatic Gaze-Based Detection of Mind Wandering during Narrative Film Comprehension. *Proceedings of the 9th International Conference on Educational Data Mining*, *8*.
- Mills, C., Graesser, A., Risko, E. F., & D’Mello, S. K. (2017). Cognitive coupling during reading. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0000309>

- Mittner, M., Boekel, W., Tucker, A. M., Turner, B. M., Heathcote, A., & Forstmann, B. U. (2014). When the Brain Takes a Break: A Model-Based Analysis of Mind Wandering. *The Journal of Neuroscience*, *34*(49), 16286–16295. <https://doi.org/10.1523/JNEUROSCI.2062-14.2014>
- Mo, J., Liu, Y., Huang, H., & Ding, M. (2013). Coupling between visual alpha oscillations and default mode activity. *NeuroImage*, *68*, 112–118. <https://doi.org/10.1016/j.neuroimage.2012.11.058>
- Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*(1), 11–18. <https://doi.org/10.1037/a0031569>
- Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1532–1543. <https://doi.org/10.1111/j.1469-8986.2011.01226.x>
- Noton, D., & Stark, L. (1971a). Eye movements and visual perception. *Scientific American*, *224*(6), 34–43.
- Noton, D., & Stark, L. (1971b). Scanpaths in eye movements during pattern perception. *Science*, *171*(3968), 308–311.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107–123.
- Peacock, C. E., Hayes, T. R., & Henderson, J. M. (2019). Meaning guides attention during scene viewing, even when it is irrelevant. *Attention, Perception, & Psychophysics*, *81*(1), 20–34. <https://doi.org/10.3758/s13414-018-1607-7>
- Pelagatti, C., Binda, P., & Vannucci, M. (2018). Tracking the dynamics of mind wandering: Insights from pupillometry. *Journal of Cognition*, *1*(1).
- 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>
- Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1–56. <https://doi.org/10.1016/j.cogpsych.2005.06.001>
- Qu, W., Ge, Y., Xiong, Y., Carciofo, R., Zhao, W., & Zhang, K. (2015). The relationship between mind wandering and dangerous driving behavior among Chinese drivers. *Safety Science*, *78*, 41–48. <https://doi.org/10.1016/j.ssci.2015.04.016>
- Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1993). Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Society for Neuroscience Abstracts*, *19*, 974. Washington, DC.
- Rayner, K. (1998). Eye movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, *124*(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, *4*(3–4), SI21–SI49. <https://doi.org/10.1080/01690968908406362>
- Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science : A Journal of the American Psychological Society / APS*, *21*(9), 1300–1310. <https://doi.org/10.1177/0956797610378686>

- Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1–21. <https://doi.org/10.3758/PBR.16.1.1>
- Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology*, *65*(2), 177–206.
- Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*(6), 1045–1060.
- 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>
- Ritchie, G. (2004). *The Linguistic Analysis of Jokes*. <https://doi.org/10.4324/9780203406953>
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747–758. [https://doi.org/10.1016/S0028-3932\(97\)00015-8](https://doi.org/10.1016/S0028-3932(97)00015-8)
- Robison, M. K., Miller, A. L., & Unsworth, N. (2019). Examining the effects of probe frequency, response options, and framing within the thought-probe method. *Behavior Research Methods*, *51*(1), 398–408. <https://doi.org/10.3758/s13428-019-01212-6>
- Robison, M. K., & Unsworth, N. (2019). Pupillometry tracks fluctuations in working memory performance. *Attention, Perception, & Psychophysics*, *81*(2), 407–419. <https://doi.org/10.3758/s13414-018-1618-4>
- Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: A database and web-based tool for image annotation. *International Journal of Computer Vision*, *77*(1–3), 157–173.
- Ruxton, G. D., & Beauchamp, G. (2008). Time for some a priori thinking about post hoc testing. *Behavioral Ecology*, *19*(3), 690–693. <https://doi.org/10.1093/beheco/arn020>
- Sayette, M. A., Reichle, E. D., & Schooler, J. W. (2009). Lost in the Sauce: The Effects of Alcohol on Mind Wandering. *Psychological Science*, *20*(6), 747–752. <https://doi.org/10.1111/j.1467-9280.2009.02351.x>
- Schad, D. J., Hohenstein, S., Vasishth, S., & Kliegl, R. (2018). *How to capitalize on a priori contrasts in linear (mixed) models: A tutorial*. Retrieved from <https://arxiv.org/abs/1807.10451v2>
- Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition*, *125*(2), 179–194. <https://doi.org/10.1016/j.cognition.2012.07.004>
- Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired? *Ergonomics*, *51*(7), 982–1010. <https://doi.org/10.1080/00140130701817062>
- Schmidt, S. R. (1994). Effects of Humor on Sentence Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. <https://doi.org/10.1037/0278-7393.20.4.953>
- Schmidt, S. R. (2002). The humour effect: Differential processing and privileged retrieval. *Memory*. <https://doi.org/10.1080/09658210143000263>

- Schooler, J. W., Reichle, E. D., & Halpern, D. (2004). Zoning Out while Reading: Evidence for Dissociations between Experience and Metacognition. In D. T. Levin (Ed.), *Thinking and Seeing: Visual Metacognition in Adults and Children* (pp. 203–226). MIT Press.
- Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319–326. <https://doi.org/10.1016/j.tics.2011.05.006>
- Schotter, E. R., Tran, R., & Rayner, K. (2014). Don't believe what you read (Only Once): Comprehension is supported by regressions during reading. *Psychological Science*. <https://doi.org/10.1177/0956797614531148>
- 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*. <https://doi.org/10.3389/fpsyg.2013.00430>
- Seli, P., Carriere, J. S. A., Thomson, D. R., Cheyne, J. A., Martens, K. A. E., & Smilek, D. (2014). Restless mind, restless body. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 660–668. <https://doi.org/10.1037/a0035260>
- Seli, P., Carriere, J. S. A., Wammes, J. D., Risko, E. F., Schacter, D. L., & Smilek, D. (2018). On the Clock: Evidence for the Rapid and Strategic Modulation of Mind Wandering. *Psychological Science*. <https://doi.org/10.1177/0956797618761039>
- 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., Cheyne, J. A., Xu, M., Purdon, C., & Smilek, D. (2015). Motivation, intentionality, and mind wandering: Implications for assessments of task-unrelated thought. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1417–1425. <https://doi.org/10.1037/xlm0000116>
- Seli, P., Kane, M. J., Metzinger, T., Smallwood, J., Schacter, D. L., Maillet, D., ... Smilek, D. (2018). The Family-Resemblances Framework for Mind-Wandering Remains Well Clad. *Trends in Cognitive Sciences*, *22*(11), 959–961. <https://doi.org/10.1016/j.tics.2018.07.007>
- Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-Wandering as a Natural Kind: A Family-Resemblances View. *Trends in Cognitive Sciences*, *22*(6), 479–490. <https://doi.org/10.1016/j.tics.2018.03.010>
- Seli, P., Risko, E. F., & Smilek, D. (2016). On the Necessity of Distinguishing Between Unintentional and Intentional Mind Wandering. *Psychological Science*, *27*(5), 685–691. <https://doi.org/10.1177/0956797616634068>
- Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-Wandering With and Without Intention. *Trends in Cognitive Sciences*, *20*(8), 605–617. <https://doi.org/10.1016/j.tics.2016.05.010>
- Seli, P., Smallwood, J., Cheyne, J. A., & Smilek, D. (2015). On the relation of mind wandering and ADHD symptomatology. *Psychonomic Bulletin & Review*, *22*(3), 629–636. <https://doi.org/10.3758/s13423-014-0793-0>
- Seli, P., Wammes, J. D., Risko, E. F., & Smilek, D. (2016). On the relation between motivation and retention in educational contexts: The role of intentional and unintentional mind

- wandering. *Psychonomic Bulletin & Review*, 23(4), 1280–1287.
<https://doi.org/10.3758/s13423-015-0979-0>
- Sheridan, H., & Reichle, E. D. (2016). An analysis of the time course of lexical processing during reading. *Cognitive Science*, 40(3), 522–553.
- Smallwood, J. (2011). *Mind-wandering While Reading: Attentional Decoupling, Mindless Reading and the Cascade Model of Inattention* (Vol. 5). <https://doi.org/10.1111/j.1749-818X.2010.00263.x>
- Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process–occurrence framework for self-generated mental activity. *Psychological Bulletin*, 139(3), 519–535.
<https://doi.org/10.1037/a0030010>
- Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: The importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology*, 4.
<https://doi.org/10.3389/fpsyg.2013.00441>
- Smallwood, J., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition*, 12(3), 452–484.
- Smallwood, J., Brown, K. S., Baird, B., Mrazek, M. D., Franklin, M. S., & Schooler, J. W. (2012). Insulation for Daydreams: A Role for Tonic Norepinephrine in the Facilitation of Internally Guided Thought. *PLoS ONE*, 7(4), e33706.
<https://doi.org/10.1371/journal.pone.0033706>
- Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., ... Schooler, J. W. (2011). Pupillometric Evidence for the Decoupling of Attention from Perceptual Input during Offline Thought. *PLoS ONE*, 6(3).
<https://doi.org/10.1371/journal.pone.0018298>
- 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., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review*, 14(2), 230–236. <https://doi.org/10.3758/BF03194057>
- Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory and Cognition*.
<https://doi.org/10.3758/MC.36.6.1144>
- Smallwood, J., & Schooler, J. (2015). The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness. *Annual Review of Psychology*, 66(1), 487–518.
<https://doi.org/10.1146/annurev-psych-010814-015331>
- Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, 132(6), 946–958. <https://doi.org/10.1037/0033-2909.132.6.946>
- Smallwood, J., & Schooler, J. W. (2013). The restless mind. *Psychology of Consciousness: Theory, Research, and Practice*, 1(S), 130–149. <https://doi.org/10.1037/2326-5523.1.S.130>
- Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as indicator and embodiment of mind wandering. *Psychological Science : A Journal of the American Psychological Society / APS*. <https://doi.org/10.1177/0956797610368063>
- Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin and Review*. <https://doi.org/10.3758/s13423-010-0046-9>

- Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*. <https://doi.org/10.1037/a0016896>
- Stawarczyk, D., Cassol, H., & D'Argembeau, A. (2013). Phenomenology of future-oriented mind-wandering episodes. *Frontiers in Psychology*, *4*. <https://doi.org/10.3389/fpsyg.2013.00425>
- Stawarczyk, D., Majerus, S., Catale, C., & D'Argembeau, A. (2014). Relationships between mind-wandering and attentional control abilities in young adults and adolescents. *Acta Psychologica*, *148*, 25–36. <https://doi.org/10.1016/j.actpsy.2014.01.007>
- Steindorf, L., & Rummel, J. (2019). Do your eyes give you away? A validation study of eye-movement measures used as indicators for mindless reading. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-019-01214-4>
- Stern, J. A., Boyer, D., & Schroeder, D. (1994a). Blink rate: A possible measure of fatigue. *Human Factors*, *36*(2), 285–297.
- Stern, J. A., Boyer, D., & Schroeder, D. J. (1994b). *Blink rate as a measure of fatigue: A review*. WASHINGTON UNIV ST LOUIS MO.
- Strick, M., Holland, R. W., van Baaren, R., & Van Knippenberg, A. (2009). Humor in the Eye Tracker: Attention Capture and Distraction from Context Cues. *The Journal of General Psychology*, *137*(1), 37–48. <https://doi.org/10.1080/00221300903293055>
- Suls, J. (1972). A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis. In J. H. Goldstein (Ed.), *The psychology of humor: Theoretical perspectives and empirical issues* (1st ed., pp. 81–100).
- Suls, J. (1983). Cognitive Processes in Humor Appreciation. In P. E. McGhee & J. H. Goldstein (Eds.), *Handbook of Humor Research: Volume 1: Basic Issues* (pp. 39–57). https://doi.org/10.1007/978-1-4612-5572-7_3
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting saliency. *Journal of Vision*, *11*(5), 5–5.
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. <https://doi.org/10.3758/BF03211656>
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. <https://doi.org/10.1016/j.actpsy.2010.02.006>
- Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, *1*(1), 29. <https://doi.org/10.5334/joc.13>
- Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review*, *13*(3), 452–457. <https://doi.org/10.3758/BF03193869>
- Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, *18*(03), 321–342.
- Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(4), 601–615. <https://doi.org/10.3758/s13415-016-0417-4>
- Unsworth, N., & Robison, M. K. (2018). Tracking arousal state and mind wandering with pupillometry. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(4), 638–664. <https://doi.org/10.3758/s13415-018-0594-4>

- Uzzaman, S., & Joordens, S. (2011). The eyes know what you are thinking: Eye movements as an objective measure of mind wandering. *Consciousness and Cognition*, 20(4), 1882–1886. <https://doi.org/10.1016/j.concog.2011.09.010>
- van den Brink, R. L., Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil diameter tracks lapses of attention. *PLoS One*, 11(10).
- 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(1), 235–242. <https://doi.org/10.3758/s13428-018-1073-0>
- von der Malsburg, T., & Vasishth, S. (2013). Scanpaths reveal syntactic underspecification and reanalysis strategies. *Language and Cognitive Processes*, 28(10), 1545–1578. <https://doi.org/10.1080/01690965.2012.728232>
- Wammes, J. D., & Smilek, D. (2017). Examining the Influence of Lecture Format on Degree of Mind Wandering. *Journal of Applied Research in Memory and Cognition*, 6(2), 174–184. <https://doi.org/10.1016/j.jarmac.2017.01.015>
- Wang, Y., Toor, S. S., Gautam, R., & Henson, D. B. (2011). Blink Frequency and Duration during Perimetry and Their Relationship to Test–Retest Threshold Variability. *Investigative Ophthalmology & Visual Science*, 52(7), 4546–4550. <https://doi.org/10.1167/iovs.10-6553>
- Wegner, D. M., & Zanakos, S. (1994). Chronic thought suppression. *Journal of Personality*, 62(4), 615–640.
- Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavior 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>
- White, S. J., Warrington, K., McGowan, V., & Paterson, K. (2015). Eye movements during reading and topic scanning: Effects of word frequency. *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), 233–48. <https://doi.org/10.1037/xhp0000020>
- Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). Sun database: Large-scale scene recognition from abbey to zoo. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3485–3492. IEEE.
- Yanko, M. R., & Spalek, T. M. (2014). Driving With the Wandering Mind: The Effect That Mind-Wandering Has on Driving Performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(2), 260–269. <https://doi.org/10.1177/0018720813495280>
- Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, 16(1), 121.
- Yarbus, A. L. (1967). Eye Movements During Perception of Complex Objects. In A. L. Yarbus, *Eye Movements and Vision* (pp. 171–211). https://doi.org/10.1007/978-1-4899-5379-7_8
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459–482.

- Zelinsky, G. J., Loschky, L. C., & Dickinson, C. A. (2011). Do object refixations during scene viewing indicate rehearsal in visual working memory? *Memory & Cognition*, 39(4), 600–613. <https://doi.org/10.3758/s13421-010-0048-x>
- Zhang, H., Miller, K. F., Cleveland, R., & Cortina, K. S. (2018). How listening to music affects reading: Evidence from eye tracking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(11), 1778–1791. <https://doi.org/10.1037/xlm0000544>
- Zhang, H., Miller, K. F., Sun, X., & Cortina, K. S. (2020). Wandering eyes: Eye movements during mind wandering in video lectures. *Applied Cognitive Psychology*, acp.3632. <https://doi.org/10.1002/acp.3632>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185. <https://doi.org/10.1037/0033-2909.123.2.162>