A REAL-TIME MIND WANDERING INTERVENTION DURING READING

A Dissertation

Submitted to the Graduate School

of the University of Notre Dame

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

by

Caitlin Mills

____

Sidney D'Mello, Director

Graduate Program in Psychology

Notre Dame, Indiana

July 2016

A REAL-TIME MIND WANDERING INTERVENTION DURING READING

Abstract

by

Caitlin Mills

Mind wandering is a ubiquitous phenomenon that occurs when attention shifts from external stimuli toward internal, self-generated thoughts. In the context of reading, it has been estimated to occur approximately 20-40% of the time (Feng, D'Mello, & Graesser, 2013; Mills, D'Mello, & Kopp, 2015). Mind wandering can have a cascading detrimental effect on reading comprehension, starting with breakdowns at the perceptual level and ending with an overall impaired conceptual understanding of a text (Smallwood, 2011). However, little is known about whether interventions can help prevent these cascading negative effects. My dissertation project developed and tested a real-time intervention for mind wandering using self-explanation practices to target deep-level comprehension processes. The goal of this dissertation was to cause a breakdown in the cascading negative effects of mind wandering during reading by detecting its occurrence in real-time and intervening by strengthening an individual's

understanding of the text. An eye-gaze based mind wandering detector was used to establish whether a person is mind wandering during reading. When this was the case, the person was asked to self-explain and potentially re-read parts of the text in order to improve their conceptual understanding of what they just read. The experimental condition was compared against two yoked-control conditions: (1) a Self-Explanation control condition that received the exact same reading conditions and interventions regardless of their mind wandering and (2) a Content-Break control that received breaks from the reading that were yoked location-wise to the interventions in the experimental condition (but did not receive the self-explanations). Results indicate that mind wandering-sensitive interventions indeed promoted better long-term conceptual comprehension compared to the Self-Explanation yoked control, particularly for parts of the texts where there was no intervention. However, contrary to predictions, the experimental condition did not exhibit higher comprehension scores compared to the Content-Break condition.

CONTENTS

FIGURES

TABLES

PREFACE

This dissertation is original, unpublished work by the author, Caitlin Mills.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Dr. Sidney D'Mello. Without his guidance, encouragement, and patience, this project would not have been possible.

I would like to thank Robert Bixler and Shelby White for all their help with the creation of materials, software, and data collection for this experiment. I would also like thank Nigel Bosch for his time spent on software development, error checking, and general support. I am also grateful for the help and guidance from Dr. Myrthe Faber who helped me throughout this process.

This research was supported by the National Science Foundation (NSF) (ITR 0325428, HCC 0834847, DRL 1235958). Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

CHAPTER 1:

INTRODUCTION

1.1 Overview

Attention is a necessity for effective learning and comprehension (Fredricks, Blumenfeld, & Paris, 2004; Hidi, 1995; Szpunar, Moulton, & Schacter, 2013; Wilson & Korn, 2007). Without attention, information can be completely missed or processed shallowly, hindering the comprehension process. This is partly due to the fact that failures in attention impede constructive processes like inference generation, memory retrieval, and other self-regulated learning strategies (Bjork, Dunlosky, & Kornell, 2013; Roediger & Karpicke, 2006). Despite our increasingly refined understanding for the role of attention in comprehension processes (Fredricks et al., 2004; Posner & Rothbart, 2005; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012; Smallwood, Fishman, & Schooler, 2007; Szpunar, Moulton, et al., 2013), research on how to respond to failures in attention is still in its infancy.

Mind wandering, which is a quintessential failure of attention, has been linked to negative educational outcomes (Smallwood et al., 2007; Smallwood & Schooler, 2006). The experience of mind wandering mirrors the anecdotal experience of reading pages of a book only to realize you cannot remember what you just read. Though your eyes continued to move across the words, your comprehension halted. Mind wandering is a ubiquitous phenomenon that occurs when there is a shift in attention from external stimuli towards internal, task-unrelated thoughts (Smallwood & Schooler, 2006). An estimated 50% of your waking thoughts are consumed by mind wandering, regardless of what you are doing (Killingsworth & Gilbert, 2010), and mind wandering is proposed to be especially detrimental in educational settings (Smallwood et al., 2007). As measured by self-reports during learning activities, the occurrence of mind wandering during educational activities is the cause for some concern: it occurs as much as 40% of the time during online lectures (Risko, Buchanan, Medimorec, & Kingstone, 2013), between 20-40% of the time during reading (Feng, D'Mello, & Graesser, 2013; Mills, D'Mello, & Kopp, 2015; Smallwood, 2011), and 18% of the time during learning with a computerized interactive tutoring system (Mills, D'Mello, Bosch, & Olney, 2015).

Significant advances have been made with respect to understanding how, when, and why mind wandering episodes arise (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009; Feng, D'Mello et al., 2013; Killingsworth & Gilbert, 2010; McVay & Kane, 2009; Mills, et al., 2013; Mills et al., 2015; Risko et al., 2013; Seli, Cheyne, & Smilek, 2013). Moreover, theoretical and empirical attempts have been made to map out the detrimental consequences of mind wandering during reading comprehension, showing that mind wandering attenuates one's ability to build a mental model of the text (Smallwood, 2011; Smallwood et al., 2007). However, despite understanding the mechanisms through which mind wandering can negatively affect comprehension, very little has been done to intervene in real-time. My dissertation takes a step toward closing this gap by investigating whether the negative effects of mind wandering on reading comprehension can be mitigated by automatically responding to mind wandering episodes in real-time with a learning intervention. I developed and tested a self-explanation-based intervention to promote deeper processing immediately after mind wandering occurs. The intervention evaluated students' conceptual understanding of the text when they are mind wandering by comparing their explanations of what they just read to "ideal" answers. If they lack understanding due to inattention, the intervention encouraged them to re-read in order to strengthen their understanding before they continue reading. The goal of the intervention being combatting the negative effects of mind wandering as they arise by intervening to ensure that students do not miss out on any important information while they are 'zoned out.'

1.1.1 Background on Mind Wandering

In recent years, the incidence of mind wandering has sparked an increasing number of cross-disciplinary research efforts, spanning from mental health, neuroscience, psychology and education (Braboszcz & Delorme, 2011; Christoff et al., 2009; McVay & Kane, 2009; Risko et al., 2013; Smallwood, O'Connor, Sudbery, & Obonsawin, 2007). Part of the reason why mind wandering has received so much attention is because of its prevalence – a striking half of our everyday thoughts (Killingsworth & Gilbert, 2010). Moreover, mind wandering is also associated with some marked benefits and costs in everyday tasks. For example, our everyday lives benefit from having the capability to zone out. Research has shown the benefits of mind wandering include the ability to engage in future planning, attend to negative feelings, or find creative solutions to problems (Baird et al., 2012; Baird, Smallwood, & Schooler, 2011; Smallwood, Fitzgerald, Miles, & Phillips, 2009). For example, people who mind wander more frequently were more likely to have creative insights in a number reduction task (Tan, Zou, Chen, & Luo, 2015).

Aside from these few (yet important) exceptions, mind wandering typically has a negative relationship with performance (for reviews, see Mooneyham & Schooler, 2013; Randall, Oswald, & Beier, 2014). In a recent meta-analysis by Randall et al. (2014), the mean correlation of mind wandering and performance, corrected for measurement error variance (ρ), was -.24 (SE = .08) across 76 independent effects (or averaged effects; k) that were reported in 88 different samples. Their meta-analysis also broke the samples into two groups based on task type: less complex ($k = 42$) and more complex tasks ($k = 38$) . Less complex tasks consisted of vigilance tasks, search tasks, and choice reaction time tasks among others. More complex tasks included tests of complex cognitive abilities. Mind wandering shared a lower correlation with performance outcomes in the less complex tasks ($\rho = -.14$, SE = .03) compared to the more complex tasks ($\rho = -.32$, SE = .04). This rather consistent negative relationship between mind wandering with performance has been the motivation for more recent research focused on attenuating its occurrence.

However, before one can respond to mind wandering, it is important to have a general understanding of how and why it arises, and its subsequent implications for learning. Prominent theoretical perspectives of mind wandering provide some insight to these questions (for a review of theories see, Smallwood, 2013). According to the Process-Occurrence framework of mind wandering, there are three different mechanisms that can explain why mind wandering arises, namely the current concerns hypothesis, meta-awareness hypothesis, and executive-failure hypothesis. The current concerns hypothesis suggests it has to do with what is most relevant to us at the time. Our minds wander about things relevant to our lives because it is more beneficial to think about goals, plans, or events that are more salient than the current external stimuli. This view posits that mind wandering occurs automatically when the incentive of current concerns outweighs the value of the current task (Klinger, 1987; Klinger, Barta, & Maxeiner, 1980; Klinger & Cox, 2011). In line with this perspective, previous studies have shown that the content of mind wandering is predominantly focused on future planning and events (Baird et al., 2011), and that listing one's current concerns before reading induces more mind wandering compared to listing non-future related knowledge based information (Kopp, D'Mello, & Mills, 2015).

Instead of current concerns, the meta-awareness hypothesis suggests mind wandering arises because of breakdowns in our meta-awareness processes (Schooler et al., 2011). Meta-representation of awareness (meta-awareness) is the ability to reconstruct and monitor the contents of our consciousness. When our attention would otherwise start to shift to a distraction that deviates from the current goal, meta-awareness monitoring can help redirect our attention back to the desired task. However, a lack of meta-awareness will lead to mind wandering because there is no longer a diversion to stop it from happening.

The third hypothesis, known as the executive-failure hypothesis, posits that mind wandering is the result of an attentional failure instead of a lapse in meta-awareness (McVay & Kane, 2009). Task-unrelated thoughts are internal distractions that are proposed to be managed by the executive control system, which typically helps focus attention while blocking out distractors (Kane & Engle, 2003). Failure to suppress task-unrelated thoughts leads to a shift in focus towards the internal distraction. Some of the main evidence for this theory comes from studies showing that people who scored lower on attentional control measures, like working memory capacity tests, tended to mind wander more and make more errors (McVay & Kane, 2009, 2012).

While the executive-failure hypothesis claims that executive control helps block out internal distractors (e.g., self-generated content) to avoid mind wandering, Smallwood (2013) suggests that executive control also plays a key role in maintaining episodes of mind wandering once they begin. This hypothesis, called the attentional decoupling hypothesis is the only one that proposes how mind wandering episodes are maintained compared to how they arise. Once attention has decoupled from external stimuli, executive control becomes responsible for the continuity of the internal train of thought that happens during mind wandering. The same mechanisms of focus and blocking out distractions are now applied to self-generated thought – where the distractors are now the external environment.

1.1.2 The Effect of Mind Wandering on Reading Comprehension

There is one commonality across all theoretical perspectives of mind wandering, regardless of how or why it occurs: mind wandering attenuates processing of external stimuli at the most basic levels of processing (Kam et al., 2011; Schad, Nuthmann, & Engbert, 2012). Smallwood and colleagues (2011) explained this phenomenon with a cascade model of inattention. Their model proposes that mind wandering impedes cognitive processing early in the perceptual stages of processing. A domino effect occurs: the effect of deficient perceptual processing cascades through the hierarchical cognitive system and ends with decoupling at higher levels of comprehension. The result winds up being sub-optimal integration of information into memory structures.

The cascading effect begins when mind wandering interferes with the basic perceptual levels of reading. Evidence that mind wandering can cause breakdowns at the level of lexical encoding while reading comes from a study by Reichle and colleagues (2010). In their study, participants' eye movements were tracked while reading Sense and Sensibility by Jane Austen one page a time. In line with models of reading comprehension (Cirilo & Foss, 1980; Graesser, Hoffman, & Clark, 1980), eye movements were sensitive to the lexical properties (i.e. word frequency) of a text when participants were on-task reading. For example, more time was spent on less frequent words. However, there was a marked difference when participants were off-task, such that their eye movements were less dependent on the lexical properties of the word. Another study echoed these findings when they found that participants' word-by-word reading times were not sensitive to lexical properties during episodes of mind wandering (Franklin, Smallwood, & Schooler, 2011). Together, these findings demonstrate that text comprehension breaks down at the early stages of reading during mind wandering. As noted above, a domino effect might occur from there, with impaired processing at the basic level of encoding leading to poor integration of information at the conceptual level.

There is also evidence suggesting a detrimental influence of mind wandering on deeper comprehension levels. While reading, individuals build a conceptual mental representation, called a situation model. The situation model guides their understanding throughout a text by keeping track of the current state of affairs and making predictions about what will happen next (Zwaan & Radvansky, 1998; Zwaan, Magliano, & Graesser, 1995). In line with the cascade model of inattention, when important information is missed due to mind wandering at basic levels, the completeness of the situation model suffers later on. For example, one study explored what happens when readers were mind wandering while critical information was given in a Sherlock Holmes story (e.g., the villain was wearing a hood; Smallwood, McSpadden, & Schooler, 2007). Participants who were mind wandering presumably missed that information. As a result, they were less likely to be able to make important inferences later in the story (e.g., John was wearing a hood, thus John is the villain) and performed worse on a memory test.

Another study provided evidence that mind wandering interferes with very basic low-level processes (sensitivity to lexical features) and cascades to high-level processes (comprehension scores) involved in reading comprehension. Mills, Graesser, Risko, & D'Mello (under review) recently tied these links together using four datasets to demonstrate a causal link between mind wandering, cognitively coupled reading strategies (low-level processes), and comprehension (high-level processes). Cognitive coupling was measured as the alignment between participants' reading times and text difficulty: reading times should be longer when the text is more difficult. They used a bootstrapping method to show that mind wandering had an indirect effect on comprehension through cognitive coupling. Results supported the cascade model of inattention (Smallwood, 2011) by showing that mind wandering leads to a reduction in coupling between the readers' efforts and the features of the text, resulting in impaired comprehension.

Taken together, theoretical and empirical accounts of mind wandering during educationally relevant tasks, such as reading comprehension, suggest the following: 1) mind wandering occurs frequently, 2) beginning at the basic perceptual level, it has a cascading negative effect on performance, and 3) is especially detrimental during more complex tasks, like reading comprehension. Together, these observations put forward the question of whether interventions targeted at restraining the cascading negative effect of mind wandering can promote better reading comprehension.

There are two important aspects to answering this question. First, interventions should promote deeper understanding in order to by-pass the negative effects of the cascade model of inattention. Improving shallow, factual memory for a text may not actually benefit overall conceptual knowledge. To this end, I focused on developing an intervention that was specifically designed to elicit conceptual understanding during reading. Second, the interventions need to be timely: they need to occur when a student has a lapse in attention. The timeliness of the intervention is important because the cascading effect of inattention posits that when mind wandering undermines conceptual understanding early on, it can have a domino effect later. For this reason, I developed an intervention that automatically responded to mind wandering at the time it arises, so that participants could update their situation model accurately before reading on. Thus, the central question of my dissertation project is: Can a real-time intervention promote deeper processing and stop the negative cascading effect of mind wandering on comprehension?

1.2 Types of Interventions

Two different methods can be used to combat the negative influence of mind wandering: (1) proactive interventions, which try to stop the cascading effect of mind wandering from ever occurring (e.g., Jha, Krompinger, & Baime, 2007; Kopp, Bixler, & D'Mello, 2014; Szpunar, Moulton, et al., 2013) and (2) reactive interventions that assuage the cascading effects of mind wandering after it has already occurred (e.g., D'Mello, Olney, Williams, & Hays, 2012). Both methods are viable solutions to minimize the negative effects of mind wandering, but each takes a distinct approach, as reviewed below.

1.2.1 Proactive Intervention

A first type of intervention attempts to reduce the negative impact of mind wandering by reducing its overall frequency through increasing one's attentional control: to stop mind wandering before it ever happens. One study sought to investigate an initial form of this type of intervention by predicting the "optimal" reading condition for each student based on individual differences (Kopp, Bixler & D'Mello, 2014). Participants' mind wandering rates were attained via auditory thought probes during reading in a study that had a 2 (text difficulty) × 2 (text value) within-subjects manipulation (Mills, Kopp, D'Mello, 2015). Texts were either difficult or easy and were manipulated as having either high or low value with respect to its weight on a subsequent test. In total, participants read four different texts on research methods, one of each combination of text difficulty and text value. Kopp et al. (2014) then attempted to predict which one of the reading conditions produced the lowest rates of mind wandering for each student using supervised learning methods. Each model was built on data from N-1 participants and then was applied to the participant that was held out. Using individual differences, they were able to predict which reading condition had lowest rate of mind wandering correctly 64% of the time (expected accuracy = 53%). Indeed, their results were some of the first to show that it may be possible to proactively reduce mind wandering by being sensitive to learning conditions.

Another proactive intervention includes the use of interpolated testing during video lectures (Szpunar, Khan, & Schacter, 2013). In this study, an interpolated test group took short quizzes after each section of the video lecture, whereas two other groups were asked to either re-study or do nothing. Their results indicated that the interpolated tests reduced mind wandering rates quite effectively. While the re-study and no-testing groups reported mind wandering 40% of the time, the tested group reported it just 20% of the time.

Most other proactive intervention attempts have been limited to mindfulness training. Mindfulness training is a non-judgmental focus on the present moment (Jha et al., 2007; Mrazek, Franklin, Phillips, Baird, & Schooler, 2013) that encourages attentional control through practicing attention regulation. When something besides the present moment comes to into focus, one should reorient their attention back to the present moment immediately. Indeed, mindfulness training has proven quite successful in several domains (Brown & Ryan, 2003; Davidson et al., 2003; Grossman, Niemann, Schmidt, & Walach, 2004) and most recently, as a way to reduce mind wandering (Ju & Lien, 2016; Mrazek, Smallwood, & Schooler, 2012; Zeidan, Johnson, Diamond, David, & Goolkasian, 2010). Previous studies provided initial evidence that greater mindfulness may reduce episodes of mind wandering (Cheyne, Solman, Carriere, & Smilek, 2009; Cheyne, Carriere, & Smilek, 2006; Mrazek et al., 2012). In these studies, people were asked to report their dispositional mindfulness using the Mindfulness Awareness Attention Scale (MAAS; Brown & Ryan, 2003), which measures how much one attends to the present moment without being distracted. As expected, results indicated a negative correlation between dispositional mindfulness and mind wandering.

Other studies then experimentally tested the effects of mindfulness training on mind wandering (Ju & Lien, 2016; Mrazek et al., 2013, 2012). Following the conventional paradigm of a longer-term mindfulness training exercise, Mrazek and colleagues (2013) had participants complete a two-week mindfulness training to test its effects on GRE reading-comprehension scores, working memory capacity (WMC) and the occurrence of mind wandering. A control condition completed sessions to learn about nutrition in lieu of the mindfulness training. Both conditions took a WMC test and a verbal GRE test before and after their training programs, but only the participants in the mindfulness condition scored significantly higher on the WMC measure and GRE after training. They also reported significantly less mind wandering in the mindfulness training condition, showing the benefits of long-term mindfulness training programs on mind wandering.

Although the Mrazek et al. (2013) study used a two-week mindfulness intervention, other training paradigms can take up to months or even years to complete and often require a meditation expert to guide the practice (MacLean et al., 2010; Speca, Carlson, Goodey, & Angen, 2000). With this demanding time concern in mind, a few other studies have explored brief mindfulness training exercises (less than thirty minutes) to reduce mind wandering (Ju & Lien, 2016; Mrazek et al., 2012). For example, after only eight minutes of a mindful breathing exercise, one study found improvement in behavioral markers that are commonly associated with mind wandering, such as fewer errors and less variability in response times in a sustained attention task (Mrazek et al., 2012). Although this study provides some promise for using mindfulness as an intervention for mind wandering, the time and resources required for mindfulness training are not readily scalable to a classroom at this point, as it is not clear how often the exercises would need to be administered or how they would influence mind wandering over time.

1.2.2 Reactive Interventions

The mindfulness interventions discussed above are proactive, as they aim to stop mind wandering from ever initially occurring. But another, equally important, question is how to respond to mind wandering when it does inevitably occur. Intervening at this level is important because we cannot completely prevent mind wandering from occurring. However, reactive interventions have been much less explored. Even in the studies that used mindfulness interventions, mind wandering rates were not reduced to zero (Mrazek et al., 2013, 2012). Additionally, it is worthwhile to consider that comprehension can suffer from any single episode of mind wandering because crucial information can be missed. Reactive responses to mind wandering offer a solution that proactive interventions cannot address by simply increasing attentional control. Instead, an intervention focused on responding to mind wandering will address potential gaps in readers’ situation models immediately as they occur. This approach aims to prevent deficits in learning caused by mind wandering, which might be particularly helpful when learners fail to realize they are mind wandering until it is too late. Responding to mind wandering episodes in real-time involves two distinct, yet critically important parts: detection and intervention, the latter being the focus on this project. In the following sections, I will review the literature on mind wandering detection followed by a discussion of previous attempts to respond to inattention in real-time learning situations.

***Mind wandering detection*.** Before you can intervene when someone is mind wandering, you have to know when mind wandering occurs. This is a difficult problem to solve since mind wandering is an elusive, internal cognitive state with few physical markers. People, including educators, often struggle to determine whether someone is completely zoned out or intently engaged. Nevertheless, recent strides have been made in detecting mind wandering while people are reading using a variety of physical markers and machine learning methods (Bixler & D'Mello, 2014; Blanchard, Bixler, Joyce, & D'Mello, 2014; Drummond & Litman, 2010; Mills & D'Mello, 2015). Below I review some of the studies that have been successful on this front, the methods they used, as well as some of the difficulties that must be overcome in order to apply a person-independent mind wandering detector in real-time.

Most of the studies discussed below used supervised machine learning models (for a review, see Kotsiantis, 2007) for mind wandering detection. This method consists of having labeled training data (e.g., mind wandering reports) and a set of features (e.g., eye gaze, heart rate, reading time, etc. leading up to each mind wandering report). For example, eye gaze features may consist of gaze duration (amount of time spent fixating on a page or screen of text) or the number of saccades that took place during the eight seconds preceding the mind wandering report. An algorithm is applied with the goal of building a supervised machine learning model that learns the relationships between the labeled training data (mind wandering or not) and the features (i.e., finding ways in which eye gaze features differentiate mind wandering from not mind wandering). Once this step is complete, the model uses the feature relationships that it has learned from the training data to test how it will perform using the features and labeled data of unseen testing data in predicting mind wandering.

The detector's performance is then evaluated based on how well it predicts the occurrence of mind wandering on the unseen testing data. Some of the common metrics that have been used to evaluate mind wandering detectors include accuracy and Cohen's kappa (Cohen, 1960). Accuracy is simply the percent correctly classified (ranging from 0 to 100%). On the other hand, Cohen's kappa takes base rates of accuracy into consideration, which is particularly helpful when classes are imbalanced. Kappa indicates the degree to which a model is better than chance at correctly predicting whether an instance was mind wandering. The kappa value is calculated using the formula κ = (observed accuracy - expected accuracy) ÷ (1 - expected accuracy), where observed accuracy is the recognition rate and expected accuracy is calculated from the confusion matrix. A kappa of 0 indicates chance, while a kappa of 1 indicates perfect agreement.

One of the first studies to attempt mind wandering detection was conducted by Drummond & Litman (2010). In their study, students were asked to read a paragraph about biology aloud and then perform a learning task (either self-explanation or paraphrasing). Periodically, they were asked to report how frequently they experienced off-task thoughts on a scale from 1 (all the time) to 7 (not at all). Students' responses were split into two categories, where 1-3 on the scale was "high" in zoning out, and 5-7 was "low" in zoning out. A supervised machine learning model was trained on the acoustic-prosodic features to classify low and high zone outs and achieved an accuracy of 64% in discriminating between the two.

This study was pivotal in automatically detecting zoning out in a learning context. However, it is unclear whether their detector would generalize to new participants. The machine learning approach that was employed used a leave-one-instance-out cross validation method instead of a leave-one-participant-out (or leave several-participants-out method). Using a leave-one-instance-out cross validation method with within-subjects data means that data from the same participant can be included in both the training and testing set. The lack of independence across participants' data in the training and testing sets can lead to overfitting: training and testing on data from the same participant means that the feature relations that the model has learned might be (too) specific to the individual participant, leading to excellent classification performance for data from that one participant, but to poor generalizability to new participants. On the other hand, participant-independent detectors are particularly useful for the purpose of real-time interventions because they need to make accurate predictions for new people, given no information about them at the outset of the learning task. For this reason, more recent detectors of mind wandering have adopted person-independent cross validation methods to ensure generalizability to new people (Bixler, Blanchard, Garrison, & D'Mello, 2015; Bixler & D'Mello, 2014, 2015; Blanchard et al., 2014; Mills & D'Mello, 2015).

Going beyond the acoustic-prosodic features used by Litman & Drummond (2010), more recent research explored different data sources collected during self-paced reading tasks (Bixler et al., 2015; Blanchard et al., 2014; Franklin et al., 2011; Mills & D'Mello, 2015). In these studies, mind wandering was measured via thought probes that occurred on pseudo-random screens (a screen of text is similar to a page of text). Participants responded either "yes" or "no" about whether they were mind wandering when the probe was presented. Binary classifiers were then trained to classify whether a person responded yes or no to the probes. Blanchard et al. (2014) used physiological features (i.e., skin conductance and temperature) to accurately classify mind wandering 74% of the time (Cohen's kappa = .22). Another study by Mills & D'Mello (2015) attempted to build a detector based on information that is readily available in log files collected during reading (i.e., reading time, textual features, etc.). They were able to classify mind wandering with an accuracy of 63% (Cohen's kappa = .21).

Franklin and colleagues (2011) achieved similar accuracy rates when they attempted mind wandering detection using a set of researcher-defined thresholds applied to log file data. This study classified if readers were "mindlessly reading" using two criterion: (1) difficulty and (2) reading time. Participants read one word on a screen at a time. Using a running window of 10 words, a specific threshold (based on pilot data) was applied to determine when readers were reading either too fast or too slow. Although this study demonstrated impressive accuracy rates for detecting mind wandering (accuracy = 72%), its methods and predetermined thresholds for fast and slow reading may not be generalizable to other, more natural, reading contexts. A number of pre-set thresholds were used with little information on how these thresholds were established, thereby complicating attempts to replicate these results or implement them for real-time interventions. Additionally, mind wandering was never predicted to occur during 'easy' portions of the text, which may not accurately reflect the real-life occurrence of this phenomenon. For example, mind wandering still occurs around 20% during easy texts (Mills et al., 2013, 2015), even though it is more frequent during difficult texts.

Perhaps some of the most promising mind wandering detection has been done with eye gaze data. Previous studies have demonstrated the feasibility of using eye gaze to accurately classify if someone is mind wandering on a screen (i.e., computer screen with text) or not, achieving accuracy rates above 70% (Bixler & D'Mello, 2014, 2015). Mind wandering detection has been successful for both probe-caught mind wandering as well as self-caught mind wandering (i.e., once the participant becomes aware that they are mind wandering) in separate studies. Each type of mind wandering report (probe and self-caught) was also found to have distinct predictive eye gaze patterns. Characteristics of probe caught mind wandering episodes included more words being skipped, longer reading times, and larger variability in pupil diameter. For self-caught mind wandering episodes, characteristics of gaze features included more regression fixations (i.e. a regressive saccade back towards text already seen) and longer saccades that crossed lines of text more often. The predicted rates of mind wandering also shared a correlation with two important variables: a positive correlation with actual rates of mind wandering and a negative correlation with comprehension (Bixler & D'Mello, 2015). Finally, these detectors were built using a leave-several-subjects-out cross-validation technique, so we can be confident they will generalize to new participants. The accuracy rates of eye gaze based mind wandering detectors will likely continue to improve as the interest in mind wandering grows, but for now, detection has reached a point where it can be applied to test real-time interventions of mind wandering.

Attempts at Real-time Interventions. Interventions and real-time feedback based on eye gaze models are becoming increasingly popular as eye-trackers become cheaper and more common. For example, eye gaze is used to increase user control during video games (Smith & Graham, 2006), improved interaction with artificial agents (Vertegaal, Slagter, Van der Veer, & Nijholt, 2001), as well as guide interventions for people on the Autism Disorder Spectrum (Wang et al., 2015). These examples highlight the reliability of using interventions with eye gaze models in real-time, albeit mostly outside of educational applications. However, a few notable studies have used eye gaze in the context of interventions for improving learning.

One study attempted to improve remedial reading ability by using eye gaze to detect when readers were struggling with recognition and pronunciation of words (Sibert, Gokturk, & Lavine, 2000). When the system detected a reader was having trouble, an auditory prompt would help students with the problematic words by giving hints. They found that readers showed improvement in reading speed and reduced errors after using their system for multiple trials. However, their sample size was quite small (N = 8) and their lack of a control condition makes their effects difficult to interpret. More recently, a study conducted by D'Mello et al. (2012) used a commercial grade eye tracker to re-engage students while they were learning from an interactive computerized tutoring system. The tutoring system, called GuruTutor, consisted of a conversation between the student and an animated agent about a biology topic. In a gaze-reactive condition, the tutor would use dialog moves to direct the student to reorient their attention when student was detected as disengaged. A control group did not receive any reactive dialog when they were disengaged. The results suggested that the gaze-reactive interventions successfully promoted learning gains for deep reasoning questions. Students in the gaze-reactive condition were also more likely to look at the important areas of screen where critical information was displayed.

These studies demonstrate the promise for improving learning using eye gaze-based detectors of students who need help or are disengaged, but only one known study by our lab has attempted to automatically respond to mind wandering using these types of detectors (D'Mello, Kopp, Bixler, & Bosch, 2016). In this study, mind wandering was detected via eye gaze while students were reading. In an experimental condition, students were asked to answer a shallow (fact-based) multiple-choice question when mind wandering was detected. If they got the question wrong, they were given a chance to re-read the text. Then, they were either asked to try the same question again or answer a different shallow question. A yoked-control condition received the same exact interventions on the same screens, regardless of whether they were detected through eye gaze to be mind wandering or not. All participants took a multiple-choice posttest after reading that tested their text comprehension on a shallow level. A median split was used to determine if participants had a high or low probability of mind wandering on a given screen based on the mind wandering detector. Preliminary analyses suggest that intervening with a shallow-level comprehension question can potentially mitigate some of the negative effects of mind wandering. When the intervention condition had high mind wandering probability and the control group's probability of mind wandering was low, there were no differences on the posttest. However, the intervention condition outperformed the control condition in one case: when the intervention condition had a low probability of mind wandering and the control condition had a high probability, scores on the posttest question were significantly higher.

This study was pivotal by showing that it is possible to respond to mind wandering in real-time, but leaves room for improvement. The intervention in this study targeted shallow processing strategies and some participants even reported "skimming" the text to look for the correct answers while they were re-reading. Since the cascade model of inattention posits missing information during mind wandering can weaken one's overall situation model, simply recalling shallow, factual information may not be the most robust way to promote deeper conceptual understanding.

1.3 Strategies to Promote Deep Comprehension

This dissertation attempted to go beyond the previous investigation by responding to mind wandering with deeper learning strategies. The goal was to promote comprehension by providing an opportunity to fill in, or strengthen, information missed during mind wandering episodes.

Proactive interventions attempt to prevent mind wandering from occurring. In contrast, reactive interventions may actually remedy the negative effects of a mind wandering by targeting the root of the problem: weaknesses or gaps in the situation model that stem from mind wandering. One way to address this problem is to draw directly from the literature on effective learning strategies. In particular, self-explanation as an effective learning strategy is unequivocal. Self-explanation involves explaining the meaning of some information to oneself during reading. According to McNamara (2004), self-explanation is a cognitively demanding, constructive activity that requires learners to actively engage in learning by meaningfully processing the material. It involves comprehension monitoring, inference generation across parts of the material, as well as integration with prior knowledge. Thus, information that is missed during mind wandering can be effectively reconstructed to create a stronger model of the text by exercising the cognitive mechanisms involved in self-explanations.

Ample evidence has shown that self-explanation is more effective for learning in comparison to many other strategies, such as re-reading and directed instruction (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Kolić-Vrhovec, Bajšanski, & Rončević Zubković, 2011; McNamara, 2004; Roy & Chi, 2005; VanLehn, Jones, & Chi, 1992). A recent study also explored which elements of self-explanation are the most effective for reading comprehension (Kolić-Vrhovec et al., 2011). This study compared various processes involved in self-explanation including elaboration, self-questioning, summarizing, paraphrasing and bridging inferences. Results indicated that elaboration (making connections across multiple parts of the text, etc.) and summarizing shared the strongest relationship with text comprehension. The authors of this study suggest these two strategies may be effective because they both require multiple levels of processing in order to organize, synthesize, and make connections across different parts of a text. Since mind wandering has a detrimental effect on processing at multiple levels, an intervention requiring exactly that type of processing could support benefits in comprehension.

1.4 Overview of Proposed Research

Mind wandering is a common phenomenon during reading that has a cascading negative effect on comprehension processes. While proactive interventions like mindfulness training can reduce the overall frequency of mind wandering, little is known about how to respond once it inevitably occurs. Can a reactive intervention stop the negative effects of mind wandering before they cascade any further? I attempted to answer this question in my dissertation by developing an intervention based on self-explanations and then testing the intervention.

According to the cascade model of inattention, mind wandering ultimately causes breakdowns in conceptual understanding throughout reading, resulting in impaired comprehension. Therefore, this project developed an intervention specifically designed to repair the breakdown before any more reading occurs – at the time when mind wandering actually happens. An eye-gaze-based mind wandering detector was used for real-time detection of mind wandering. When mind wandering was detected, participants were prompted to engage in self-explanations before moving on in order to boost participants' conceptual understanding. As mentioned above, self-explanations can help readers diagnose their gaps in understanding, help them elaborate on ideas, and make inferences to strengthen their overall understanding of a concept (Bjork et al., 2013; Kolić-Vrhovec et al., 2011; McNamara, 2004). Thus, the intervention will try to ensure that readers have mastered a concept before they simply move on.

CHAPTER 2:

DEVELOPING A SELF-EXPLANATION INTERVENTION

2.1 Overview

A successful intervention would mitigate the negative effects of mind wandering (proposed by the cascade model of inattention). With this goal in mind, there were four considerations that needed to be addressed when designing the intervention: which text to use, how to detect mind wandering, how to implement self-explanations, and how to score participants' answers. Each is discussed in detail below.

2.2 Text

The text was a 35-page excerpt from a book on a mundane topic that would likely be unfamiliar to most participants. The excerpt was taken from a book entitled "Soap-bubbles and the Forces which Mould Them" (Boys, 1895). The excerpt contained approximately 6,535 words taken from the first 35 pages of the book, which were then broken down into 57 screens of text. The original text was accompanied by images and some references to the images which were removed from the text. Removal was done to ensure comprehension would not be impaired by the absence of the images or explanations. Beyond these modifications, the text was consistent with the original published version. The text can be found in Appendix A.

There were approximately 115 words of text on each screen. This number was chosen because it allowed for a relatively ecological reading task with multiple paragraphs on a screen at once. At the same time, restricting the words per screen to 100 allowed for the lines to be spaced far enough for more accurate eye-tracking. An example of a screen is included in Appendix B.

For purposes of the intervention and knowledge assessments, the 57 screens were also broken down into fifteen sections based on distinctive concepts in the text. For example, the first three screens are dedicated to explaining the authors' motivation for writing the book, whereas a later section explains an experiment on how a water sieve works. All of the sections were between three and five screens long to represent fifteen different concepts. See Appendix C for a breakdown of the sections, screen numbers, and target concept.

2.3 Mind Wandering Detector

The detector used in this study was built using a previous dataset and its validation, accuracy and implementation are discussed below. The validation is necessary in order to have some confidence that mind wandering can be reliably estimated, despite having no background information on a given participant at the time they begin the study.

2.3.1 Validation

The mind wandering detector was originally built using a dataset where participants read the same exact text as described above (for a description of the original dataset used, see Kopp et al., 2015). Participants self-reported mind wandering during reading whenever they caught themselves doing so (self-caught reports). Each screen with a self-report was used as an instance of mind wandering, while each screen without a self-report was used as an instance of normal reading. The model was evaluated based on how accurately it could distinguish between the two types of instances.

An eye tracker provided gaze data for each eye including the x and y position of the gaze on the screen, and the pupil diameter. This stream of data was converted into a series of fixations (periods where the eye remained fixated on the same point), saccades (rapid eye movements between fixations), and blinks, from which features were calculated. In general, the term feature refers to variables that are created or observed in order to predict the dependent variable of interest. The mind wandering detector described here uses a set of global eye gaze features, which are independent of the words being read. Features were calculated from only a certain range of gaze data from each screen, called the window. Varying window sizes (i.e., duration of time in which features were calculated from) were tested (4, 8, and 12 second windows) and compared based on their performance. For screens without a self-report, the end point of the window was positioned 16 seconds into the screen, which was the average number of seconds between a self-report and the beginning of the screen. For pages with a self-report, the end point of the window was positioned 3 seconds prior to the self-report in order to avoid picking up on changes in the gaze data associated with reporting mind wandering (i.e., looking down at the keyboard to press a button). Screens that were shorter than the window size were discarded, as were screens with windows that contained fewer than five fixations (i.e., not enough data to make a prediction).

2.3.2 Accuracy

Supervised machine learning algorithms available in the Waikato Environment for Knowledge Analysis (WEKA; Hall, Frank, Holmes, Pfahringer, Reutemann & Witten, 2009), a machine learning software "workbench", were used to build models which discriminated instances of mind wandering (screens with a self-report) from instances of normal reading (screens without a self-report). A variety of classification algorithms, including Bayesian models, support vector machines, rule-based classifiers, regression models, lazy learners (such as K-nearest neighbors), and decision trees were tested. The models were validated using a leave-one-participant-out cross-validation method in order to ensure that data from each participant was exclusive to the training or testing set. This entailed training the model on data from n-1 participants and testing the model on the remaining participant a total of n times. The kappa value was used to evaluate the detector's performance, which was computed by combining all of the classifications from each participant into a single confusion matrix. A support vector machine (SVM) using an 8-second window was the best performing model, attaining a kappa value of .22 (22% more accurate than chance level).

2.4 Detector Implementation

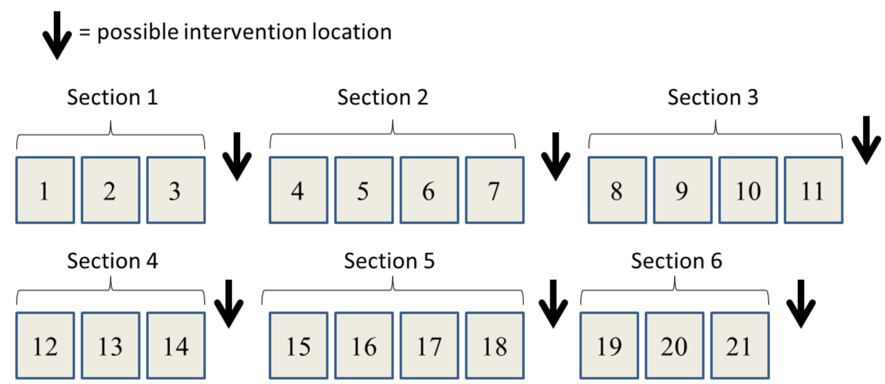
As part of my dissertation research, the validated detector was then implemented for use in a real-time system. Eye gaze features were computed automatically as participants read which were then used by the SVM to predict mind wandering on every screen. Parameters of the SVM were set based on the final model produced from the training data discussed in the validation section.

Implementing the detector also required establishing the criterion by which an intervention will be deployed based on the detector's likelihood of mind wandering output. On every screen, the detector outputs a likelihood that the screen may result in a positive instance of mind wandering, expressed as a number between 0 and 1. Interventions only occurred at the end of a conceptual section, at which point the average mind wandering likelihood value from that section was used to make a decision about whether an intervention should be initiated. Figure 2.1 shows a graphical example of how the interventions were triggered. The second conceptual section in the text contains screens 4-6, so three likelihood values from these screens were averaged to make the decision. The intervention appeared when the participant tried to advance from 6. The third conceptual section contains five screens (7-11), so five values were averaged to make a decision on screen 11. The average was chosen because it takes into account all the screens participants would need to see in order to understand a concept.

In addition, initial pilot data (N interventions = 59) revealed the average mind wandering likelihood shares a small negative correlation with participants' self-explanation scores ($r = -.12$).

The decision likelihood value is then used to decide whether an intervention should occur. Rather than using a single pre-set threshold to make a binary decision, this project used a probabilistic method to modify the chances of an intervention based on the average likelihood value. This is important because the mind wandering detector itself is imperfect. At this point, there is no absolute certainty for when mind wandering will occur, and no perfect way to implement an accurate binary distinction in the likelihood values. Thus, using a probabilistic function allows for an informed application, without effectively reducing the amount of interventions by limiting their occurrence to 100% confidence.

The probabilistic decision method was implemented using a step-wise function (see Appendix D). At each conceptual section, an intervention was never deployed if the average mind wandering likelihood falls along the bottom of the curve (anywhere below .30). If the likelihood falls in the middle section from (.30 - .66), the participant's average likelihood score was used to probabilistically determine if they will receive an intervention. For example, if a participant had an average likelihood of .40 for a given section, they had a 40% chance of getting an intervention. Finally, if the average mind wandering likelihood is above .66, participants always received an intervention for that section.



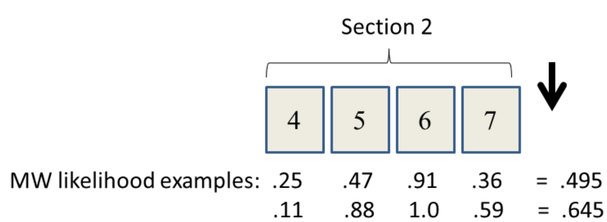


Figure 2.1: Example of how screens are broken up into sections, where interventions can come, and how the average mind wandering likelihood is computed per section (the average of the section's likelihood scores for two participants).

These cutoff points were decided after examining histograms and scatterplots from pilot data. That is, at least 25% of pilot data fell within the never receive section of the function and 25% fell within the always receive an intervention portion of the curve. Moreover, the step-wise function method was adopted instead of a strictly linear probabilistic approach because the values on either end of the average mind wandering likelihood distribution should not be as probabilistic as the values in the middle, where there is less certainty. For example, an average likelihood of .21 was the minimum likelihood seen in the pilot data and should be given an intervention instead of having a

21% chance of moving on without an intervention. Using an even more conservative threshold would not be ideal because the average mind wandering likelihoods will be unlikely to fall in that range in the first place.

2.5 Self-Explanation Questions

The goal of the intervention was to ensure that the reader understands individual concepts as they arise (one concept per section of text) before moving on to the next one. An intervention could have occurred whenever there is a break in concept unit in the text (at the end of each of the 15 demarcated concept sections). These interventions targeted some of the mechanisms described in McNamara’s (2004) definition of self-explanations like understanding gaps in understanding, explaining and elaborating on concepts, and tying explanations to prior knowledge. Questions asked participants to elaborate on a concept described in the text in order to demonstrate understanding. For example, participants were asked to explain what the author’s point was or how a concept from the text would apply to another real-world situation (e.g., Based on an experiment described in the text, what would happen to your clothes when you jump in a pool with them on?). The questions were previously piloted using students from Notre Dame’s subject pool to make sure they were not too easy or too difficult. See Appendix E for intervention questions.

The interventions proceeded with the following flow of events (depicted in Figure 2.2): First, mind wandering was detected. Next, an intervention was triggered. It appeared on the right side of the screen and the text that they were reading was grayed out. At this time, participants received a question that asked them to explain a concept related to that section of text. They were not able to go back and re-read the text at this time and were forced to answer based on their current understanding (first attempt). This was done so that they cannot simply use the text to formulate their answer as they type. If their self-explanation was accurate enough, they received positive feedback and moved on. If their answer was deemed incorrect, they received neutral or negative feedback and were asked to re-read the last few screens in order to provide a better answer (second attempt). Re-reading is used as a way to give participants an opportunity to go back and fill in any gaps in their understanding, which they may have realized during their first self-explanation attempt. Participants indicated they were ready to try the question again once they were done re-reading by clicking a button saying ‘Try Again’. While answering the intervention question, they were once again unable to read through the text until submitting their second response.

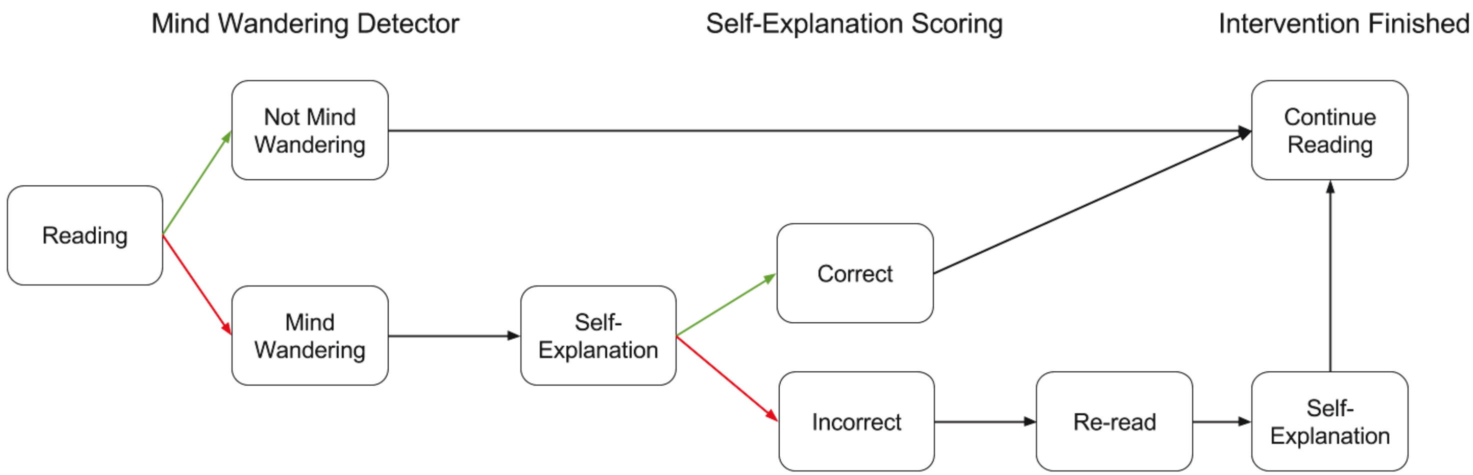


Figure 2.2: Intervention Protocol Flowchart.

2.5.1 Intervention Scoring

Feedback for the participants' self-explanations to the intervention questions also needed to occur in real-time. Since these are open-ended questions, a proper scoring technique presents one of the major challenges for the current project. After attempting a variety of approaches to scoring the students' answers, the best approach came by combining three different techniques. First, a basic word overlap algorithm was used, which matches words across two documents being compared. Participants' answers were compared to predefined "ideal answers" for each question and the overlap score was equal to the proportion of words that overlap across the documents. However, it is important to consider that participants' answers may contain synonyms that do not precisely match with ideal answers, yet are correct in their meaning. Thus, the second step was to obtain synonyms using WordNet (Fellbaum, 1998), a publicly available lexical database that contains groups of words that denote a specific concept (called a synset). Synsets were obtained for each of the words present in the ideal answers and word overlap scores included any overlap with the synonyms as well. Finally, critical words and phrases that are essential for a correct answer were identified for each of the 15 questions (up to five per question). Two human coders reviewed the ideal answers and identified keywords present in the ideal answers (e.g., capillary action). If a participants' answer contains one of these keywords, their word overlap score was boosted by .20.

Participants' answers were automatically scored by comparing them to a total of ten different ideal answers to obtain a word overlap score. The highest word overlap score was taken as the participant's score plus any key word boost they receive (up to .20 additional points). To evaluate this scoring method, automatically generated word overlap scores were compared to human ratings of the self-explanations for a set of pilot data containing 151 self-explanations. A human coder rated the self-explanations on a scale of 0 (completely incorrect) to 1 (completely correct). Then, the automatic word overlap scores were correlated with the human ratings of correctness, yielding a correlation of .667, which is acceptable based on a review of the other attempts to automatically score short-answer, open-ended responses (see Burrows, Gurevych, & Stein, 2015).

2.5.2 Intervention Implementation and Feedback

The intervention was implemented based on two decisions: if participants should be given the chance to re-read and the type of feedback they will receive. Since self-explanations were scored on a scale from 0 to 1, there is no binary cut off for overlap correctness. Therefore, similar to the mind wandering detector, a step-wise function was also used to decide if participants would re-read using three different cutoffs. Word overlap scores from 0 to .30 always required participants to re-read (re-read cutoff). For scores from .31 to .69, the inverse of participants' word overlap scores was used to probabilistically determine if they will re-read (probabilistic cutoff). For example, a 40% overlap yielded a 60% chance of re-reading. Finally, scores ranging from .70 to 1.0 were always able to move forward in the text (keep reading cutoff).

Another issue that comes along with the subjective nature of scoring self-explanations is the fact that giving participants incorrect feedback could be demotivating and cause negative, reactive behaviors. Therefore, three types of feedback were given to participants: negative, (e.g., "Not quite! Let's try again."), neutral (e.g., "Thanks for providing an answer. Let's continue."), or positive (e.g., "Great answer! Let's keep reading!"). The re-read cut offs always received negative feedback and the keep reading cut off were always given positive feedback. The middle section received positive feedback if they are able to move on, but neutral feedback if they are forced to re-read, avoiding negative feedback since answers were at least above 30% word overlap.

If participants were asked to re-read based on the word-overlap scoring algorithm, the answer box disappeared after receiving feedback and they were able to re-read, navigating back and forth using the arrow keys. When they were finished re-reading, they pressed a button on screen labelled "Try Again" in order to type in a new answer (see Appendix F for an example of the intervention screen). Once they pressed the "Try Again" button, they were unable to navigate through the text. An empty answer box reappeared for them to type in their second self-explanation attempt. Submitting their second answer concluded the intervention, regardless of whether they got it correct. Participants received feedback on their second answer and then progressed to the next screen after they were finished with the intervention questions (either after one or two attempts).

2.6 Research Questions and Design Considerations

The intervention was designed to answer three overarching research questions: (1) can a mind wandering-sensitive self-explanation intervention repair comprehension deficits in real-time, as participants are reading? (2) can this self-explanation intervention prevent the cascading negative effect of mind wandering on overall comprehension? and (3) what factors, if any, influence the effect of the intervention on comprehension?

To test whether the self-explanation interventions mitigate the effect of mind wandering, I used a between-subjects yoked-control design with three conditions: MW Sensitive condition (experimental condition; MW), Self-Explanation control (SE control), and Content-Break control (CB control). All participants read an excerpt from a scientific book on surface tension and soap bubbles and were given a final comprehension test after the reading task was complete. One group (MW condition) received learning interventions (via self-explanations) only when they were “needed” (i.e., when a gaze-based detector indicated that the participant was likely to be mind wandering). For every participant in the MW condition, data for two yoked control participants were collected. In the SE control condition, participants received an identical intervention at the same place in the text, regardless of whether they were mind wandering or not. Participants in the CB control condition also received interventions at the same place as the MW condition irrespective of their mind wandering rates, but instead of getting a self-explanation intervention, the CB control answered tip-of-the-tongue questions to provide a content break from the text. Participants were not matched based on other criterion.

The SE yoked condition provides a controlled comparison to the MW group because it provides participants with self-explanation interventions at the same points as the MW group, irrespective of whether they are mind wandering. In this way, it is possible to assess whether intervening with the self-explanation at the right time during reading in the MW condition (i.e. when the cascade effect of inattention begins) is more effective than intervening in general (i.e. when mind wandering is not occurring). Furthermore, a yoked-control was chosen over alternative designs to avoid additional confounds across conditions. For example, in a “do-nothing” control in which participants did not receive any self-explanation interventions, any advantage seen in the MW condition could be due to the self-explanations and re-reading, rather than the timely interventions. Similarly, because the self-explanation offers a periodic break from the text, and a do-nothing control would lack such breaks, the SE condition provides a tighter level of control.

Because the self-explanation interventions provide this inherent break from reading, a second yoked condition was included to control for possibility that simply giving participants periodic breaks might also improve comprehension. This idea is rooted in models of event cognition, which suggest that event boundaries can influence memory and comprehension (Pettijohn, Thompson, Tamplin, Krawietz, & Radvansky, 2016; Radvansky, 2012; Swallow, Zacks, & Abrams, 2009). When an event boundary is encountered, the mental model of the text is updated, causing information related to the situation before the boundary to be committed to longer-term memory storage. Therefore, the CB control condition provides a baseline comparison between participants who experience a break, but without the opportunity to accrue comprehension benefits of self-explanation, and the MW participants, who had such an opportunity.

The design of the SE condition also controls for potential effects of interpolated testing. Interpolated testing can be a proactive mind wandering intervention (Szpunar et al., 2013), and thus it is important that participants across these conditions engaged in the same number of self-explanations as the experimental condition. Additionally, because self-explanation has a strong record of being an effective learning strategy, it would be problematic to compare learning across the conditions if they received different numbers of interventions. For example, if the SE control participants received interventions at every conceptual break, rather than selectively (as in the MW condition), we would expect the positive effects of these self-explanations to surpass the benefits of receiving fewer interventions that are linked to mind wandering. Thus, the intervention was designed to assess whether learning is improved by triggering self-explanation interventions only when they are needed, rather than assessing the effects of self-explanation as a general learning strategy.

2.6.1 Research Question 1

The cascading negative effect of mind wandering on reading comprehension begins at a local level: at some point during the reading process, participants mind wander and thus do not effectively encode information (Smallwood, 2011). Remediation at this basic level is a necessary first step towards improving overall comprehension. Therefore, my first research goal was to attenuate the effects of mind wandering during reading at the concept level. Participants in the MW (experimental) condition received interventions when they were mind wandering, where comprehension is expected to be lower. As such, I predict that self-explanations provided by participants in the MW condition will improve on their second attempt after re-reading (within-subjects comparison). However, participants in the SE control condition were expected to exhibit no differences in self-explanation scores since the timing of the interventions was not contingent on their mind wandering. Figure 2.3 shows a graphical depiction of the predicted pattern for research question 1. The CB condition cannot be evaluated with respect to online comprehension because they did not engage in any self-explanations.

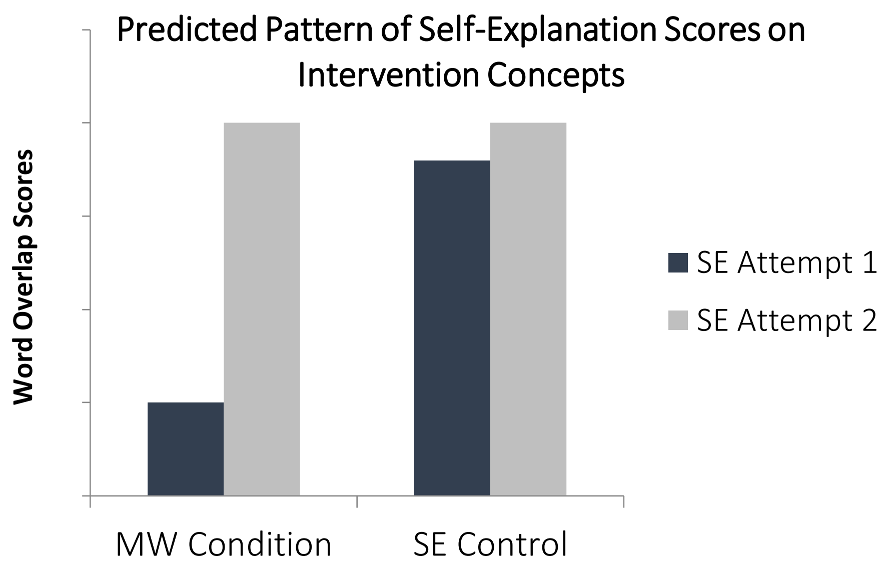


Figure 2.3: Predicted pattern for research question 1. When participants are mind wandering and get an intervention (in the experimental condition), performance will improve from their first to second self-explanation attempt.

2.6.2 Research Question 2

Research question 2 focuses on how the mind wandering sensitive intervention influenced overall comprehension, beyond its effects on immediate repairs to online comprehension. Lower comprehension was expected for concepts presented while participants were mind wandering compared to when they were paying attention (Randall et al., 2014; Smallwood, 2011). If comprehension is successfully repaired when mind wandering occurs (research question 1), then the MW intervention should lead to a better overall understanding of the text. The control conditions did not always receive interventions when they were needed, thus their overall understanding of the texts is likely to suffer. Therefore, I predict that participants in the MW condition will outperform their yoked counterparts in the SE control and the CB control conditions on the posttest comprehension measures.

For the posttest comparisons, the conditions will be compared on multiple dimensions of comprehension (four tests in total). Two levels of comprehension that are compared are text level comprehension (surface level facts from the text) and inference level comprehension (requiring deep reasoning and conceptual understanding). These two levels of comprehension will be assessed immediately after reading (immediate tests) and then again one week later (delayed tests). Although the patterns are expected to be the same across the different dimension of comprehension, it is important to examine each independently since they are conceptually distinct. Thus, immediate text level, immediate inference level, delayed text level, and delayed inference level comprehension will be analyzed separately.

2.6.3 Research Question 3

Participants in the different conditions had different experiences on the intervention concepts versus the non-intervention concepts. For example, participants in the MW condition received an intervention when they were mind wandering, while participants in the control conditions received interventions when they were less likely to be mind wandering . Conversely, participants in the MW condition were unlikely to be mind wandering on the non-intervention concepts.

Therefore, the third and final research goal is to test whether the occurrence of an intervention for a given concept moderates any overall effects of condition on comprehension. For example, do the MW condition and SE condition experience the same benefits in comprehension for intervention concepts, despite the fact that the MW participants were likely more likely to be mind wandering than their SE counterparts? How do the two groups perform on comprehension questions specific to the non-intervention concepts? These questions can be answered by analyzing comprehension questions corresponding to intervention concepts and non-intervention concepts separately.

The non-intervention concepts correspond to situations where the MW condition was not mind wandering. The MW condition is thus expected to perform better than the SE condition because they may or may not be mind wandering. However, I do not expect to find any differences in intervention-relevant comprehension between the MW and SE control condition. Although the mind wandering intervention may bring the MW condition up to par with the SE control on intervention-relevant comprehension, it is unexpected that they would perform significantly better, as it is unlikely that SE control was mind wandering and they also received a learning intervention.

The MW condition is also predicted to outperform the CB control on non-intervention comprehension since they were not predicted to be mind wandering. Unlike my prediction for the SE control, I also expect that the MW condition will outperform the CB condition on the intervention-relevant comprehension. Despite the fact that they were mind wandering, they still received self-explanation learning intervention, while the CB control condition simply received a periodic break from the reading task.

CHAPTER 3:

METHODS

3.1 Participants

Eighty-one undergraduates from a private mid-western university participated for course credit. The average age was 19.24 (SD = 1.01, 59.2% female). The majority of the participants were Caucasian (65.4%), with 14.8% Hispanic, 12.3% African American, 6.2% Asian, and 1.2% responded as “Other”.

Participants were given the option to return one week later to complete the delayed tests for additional course credit. Out of the 81 participants, 54 of them (67%) opted to come back.

3.2 Design

The experiment had a between-subjects yoked-control design in which participants read 57 screens of text separated into 15 conceptual sections about how surface tension and soap bubbles work. The three conditions were MW (mind wandering-based learning interventions), SE control (yoked self-explanation intervention), and CB control (yoked content break intervention). When our mind wandering detection model predicted that participants in the experimental (MW) condition were mind wandering, they were asked to provide a short explanation about key concepts and relationships from the text at the next concept boundary. They were only given an opportunity to re-read the text if their original answer was wrong (61% of the time).

The SE yoked control condition received interventions at the same locations as their experimental counterpart. Thus, participants in the yoked control condition provided self-explanation for the same exact questions, but the occurrence of interventions was not sensitive to their attention. Additionally, the SE condition was be given the same opportunities to re-read and re-answer the self-explanations in order to control for exposure to the text. The SE condition was given neutral feedback in the event that they had to re-read despite performing well on the word overlap.

After a participant in the experimental condition completed the experiment, the next participant was yoked to their experiences. As mentioned above, a yoked design approach allows for careful control over the frequency of interventions and their locations. This way, a distinction can be made between the benefits of giving a self-explanation learning intervention in general versus the effects of giving an intervention at the appropriate time (when mind wandering has occurred).

3.3 Materials

Eye Tracker. A Tobii TX 300 eye tracker was used to collect gaze data in real-time. Participants' eye gaze was calibrated at the beginning of the study. An experimenter examined the calibration quality before the reading task began.

Tip-of-the-Tongue Questions. A total of 40 tip-of-the-tongue questions with four multiple choice answer options as part of the content break for the CB control condition. The set of questions were found in Appendix A of Meyer & Bock (1992). See Appendix G for a list of questions.

Knowledge Assessments. Students took two sets of knowledge tests, one set immediately after completing the reading task and a second set delayed until a week after the initial session. Each set included a Text Level Comprehension test and an Inference Level Comprehension test.

Text Level Comprehension tests each consisted of 30 text level questions that required surface level knowledge (e.g., facts about the text). Sixty questions were randomly split to make two versions of the Text Level Comprehension test. The two versions were counterbalanced across participants to be presented as the immediate or delayed version of the test. The test items also corresponded to a section of the text, so comprehension scores could be tied to specific interventions. All 60 text level questions were created for a previous study and can be found in Appendix H (Kopp et al., 2015).

An Inference Level Comprehension test consisted of deep reasoning questions that target conceptual knowledge of the text (see Appendix I). A deep reasoning question is one that requires making some inference or application of knowledge in order to answer the question correctly (Graesser, Ozuru, & Sullins, 2010; Graesser & Person, 1994). Each question was designed to target conceptual knowledge of a specific section in the text (30 total, 2 per section) and was a 4 alternative, forced choice questions. The four answer options consisted of a target (the correct response to the question), a near-miss (an option that sounds correct but was not), a thematic miss (an option that follows the theme of the content but is not actually related to the question), and a distractor (an option that is not at all related).

The questions were piloted using Amazon's Mechanical Turk to ensure they were not too easy or difficult to answer. This was tested by first asking participants to answer the questions without having read the text. In this step, accuracy was expected to be around chance levels. Modifications were iteratively made at this point to ensure the question could not be answered too easily without having seen the text (i.e., if everyone gets a question correct based on prior knowledge, it was removed or modified and re-tested). The next step required participants to answer questions after reading the associated conceptual section of text. In all, there were two questions per conceptual section (30 questions total), which were randomly split to make two versions of the Inference Level Comprehension test. The test contained one question for all of the 15 concepts. The two versions were counterbalanced across participants and presented as either the immediate or delayed version of the test.

3.4 Procedure

Participants completed the experiment individually over a 1.5 hour session (plus a follow up session). The experiment occurred in four separate phases: (1) calibration and instructions, (2) reading task with interventions, (3) posttests and (4) delayed posttest.

3.4.1 Phase 1 Calibration and Instructions

Participants signed an informed consent acknowledging their rights as a participant. Participants were then seated in front of a computer in order to begin the experiment and completed the calibration process, which took about two minutes to complete. Calibration involved focusing on a series of five yellow dots that appear on the screen in five different locations, then looking at a set of numbers at different places on the screen. Once a participant has finished the calibration process, the experimenter was able to view the accuracy of their calibration on a different computer. If the calibration was accurate, participants moved on. If the calibration was inaccurate, the calibration process was repeated up to three times before moving on. Calibration was a subjective judgment made by trained experimenters, so there is no numerical threshold required to move on. Participants were not required to have uncorrected vision and the eye tracker was able to collect data even with glasses on.

Next, participants received instructions about the reading task: “Your primary task is to read the text in order to take a short test after reading. The text will be displayed screen by screen and you can press the "right arrow" key to navigate through the text and the "left arrow" key to return to the previous page.” Then, participants were given notice of the intervention questions, but they were not told how or when the questions would appear. The following instructions were be presented to participants via computerized instructions: “While reading the text, you will occasionally be asked some questions about the pages that you just read. Depending on your answer, you will have the chance to re-read and answer again. You may have to review the last few pages that you have read. Remember, you can use the left arrow key to review what you have read."

3.4.2 Phase 2: Reading Task with Interventions

The texts were displayed using size 36 Courier New font with approximately 100 words per screen (for 57 screens) on a 511 × 287 mm monitor running at 1920 by 1080 resolution. Participants were instructed to read each screen and press the right arrow when they were ready to advance to the next screen or press the left arrow to return to the previous screen.

Interventions occurred throughout the text. Examples of what the intervention screens looked like are given in Appendix F. For the MW condition, interventions were triggered when they were mind wandering. The yoked control participants received interventions on the exact same screens as their counterpart from the intervention condition. Participants typed their answers directly into a text box on the right-hand side of the screen. They clicked a button to submit their answer, which was automatically scored. Participants then received feedback based on the probabilistic feedback function. The intervention screen looked similar in the CB condition, the only exception was the self-explanation answer field was replaced with a box with multiple choice tip-of-the-tongue questions. Participants in the CB condition answered tip-of-the-tongue questions for as long as the experimental condition re-read the concept. If they were in the middle of a question when the timer went off, they would simply move on after the question was finished. Feedback was not given about the correctness of their answers to the tip-of-the-tongue questions.

No more than nine interventions were given to a participant so that they were not inundated with interruptions throughout the text. When the reading task is over, participants were instructed to get the experimenter.

3.4.3 Phase 3: Posttests

Participants filled out two different knowledge measures immediately after the learning session. First, they completed a 30-item Text Level Comprehension test, consisting of surface level questions. Next, participants answered the 15-item Inference Level Comprehension measure. Finally, participants filled out a brief demographics questionnaire and a follow-up session was scheduled if participants were willing to come back for an extra .5 research credits. Participants in the yoked control conditions received the exact same questions as the MW condition on each of the posttests.

3.4.4 Phase 4: Delayed Posttest and Debriefing

Participants were given the opportunity to come back after one week for more research credit in order to complete the delayed tests. When they came for delayed testing, participants completed alternative versions of the Text Level Comprehension posttest and Inference Level Comprehension posttest. Since there were two versions of the tests, participants received whichever version they did not receive previously. Finally, participants were fully debriefed.

CHAPTER 4:

RESULTS

Not every member of a given triplet completed the delayed test (e.g., missing either the CB control or SE control participant). Thus, analyses for the delayed tests could not be completed at the triplet level while at the same time maximizing the number of participants included. Therefore, analyses were all completed at the yoked-control pair level in order to keep the analytic approach consistent throughout the analyses. In this way, differences between the yoked-controlled pairs (MW condition vs. SE control and MW condition vs CB control) were evaluated using the same criterion for each of the four types of comprehension (immediate text level, immediate inference level, delayed text level, delayed inference level). In total, there were 27 pairs for the immediate tests and 13 pairs for the delayed tests, respectively. Furthermore, since the main research questions were focused on the effectiveness of a mind wandering sensitive learning intervention relative to two different yoked controls, the SE control and CB control conditions were not directly compared.

The variables of interest were computed as follows: for each participant, I computed the average mind wandering likelihood scores (across all 15 concepts), as well as their average word overlap score (across all intervention concepts), and comprehension scores. The four comprehension tests were scored by computing each participant's proportion of correct responses on each of the tests separately. All participants received at least one intervention and were therefore included in the analyses. Significance testing was done with an α set to 0.05.

4.1 Detector Overview

Participants received 4.89 interventions on average (SD = 2.49) across the 15 possible intervention concepts. This intervention rate of roughly 30% is consistent with self-reported of mind wandering rates during reading (20-40%). In total, there were 396 interventions across the three conditions.

I first correlated the detector's predicted likelihood of mind wandering with online comprehension (participants' first self-explanation attempt) for participants in the MW condition. If the detector was in fact picking up on episodes of mind wandering, its predictions should be negatively related to online comprehension, establishing some evidence for convergent validity. Participants' average likelihood of mind wandering on intervention concepts was negatively correlated, but not significant, with their average word overlap scores from their first attempts to answer intervention questions $r(27) = -.269$, $p = .175$.

Average mind wandering likelihood scores were expected to be negatively related to performance on the overall comprehension tests in the MW condition as well. However, the correlation was not expected to be as strong for overall comprehension tests as the real-time interventions may have mitigated the negative relationship between mind wandering and overall comprehension. Average mind wandering likelihood scores were weakly, but negatively, correlated with performance on the text-level comprehension tests, $r(27) = -.148$, $p = .460$ for immediate and $r\ (27) = -.155$, $p = .526$ for the delayed test. Negative correlations were also found between mind wandering likelihood and performance on the inference level tests, though they were also quite small; $r(27) = -.061$, $p = .762$ for immediate and $r(527) = -.102$, $p = .677$ for delayed.

4.2 Does the Intervention Improve Online Comprehension during Mind Wandering?

The first research question was whether the intervention attenuated the negative effects of mind wandering during reading in the MW condition. To answer this question, word overlap scores were compared between the first self-explanation attempt and the second attempt. Three participants scored high enough on their first word overlap attempts as to never receive a second attempt, so this analysis was completed with 24 participants.

In the MW condition, participants received an intervention when they were likely to be mind wandering, and presumably not deeply processing the text. Thus, they were expected to improve after re-reading and trying to answer again. A paired samples t-test revealed that participants' word overlap scores significantly increased from first attempt ($M = .379$; $SD = .156$) to their second attempt ($M = .513$; $SD = .151$), $t(23) = 3.03$, $p = .003$. For the SE control condition, word overlap scores also increased from the first attempt ($M = .453$; $SD = .184$) to the second attempt ($M = .500$; $SD = .152$), but this difference was not significant, $t(23) = 1.31$, $p = .202$.

The MW condition and the SE control condition were also compared on their first and second self-explanation attempts. The SE control was expected to perform better on the first attempt since they were not likely to be mind wandering, but they were expected to perform similarly on the second attempt after the chance to re-read. For their first self-explanation attempt, the MW condition ($M =.412$; $SD = .179$) did not perform significantly worse than the SE control ($M =.461$; $SD = .178$), $t(26) = 1.25$, $p = .224$, $d = .276$. Although the difference was statistically significant, the effect size ($d = .276$) is consistent with a small effect of condition (Cohen, 1992). The MW condition ($M =.513$; $SD = .151$) and SE control ($M =.500$; $SD = .152$) did not differ in their second self-explanation attempts, $t(23) = .326$, $p = .748$, $d = .086$. This suggests that their self-explanations were on par by the time they finished an intervention.

4.3 Do Mind Wandering-sensitive Interventions Improve Overall Comprehension?

A central goal of this dissertation was to determine whether participants in the MW condition would outperform the two control conditions on tests of comprehension. This question was addressed using within-subjects comparisons between the MW condition and the two yoked control conditions. The yoked control design is well-suited to this type of comparison because variability across subjects is removed with respect to three important factors: the intervention itself (e.g., having a chance to engage in self-explanation), the location and content of the intervention, and the frequency of intervention.

Repeated measures comparisons were made between the experimental condition (MW) and each of the two yoked control conditions (SE and CB). A total of eight repeated measures tests were conducted: 2 comparisons (MW condition vs. SE control and MW vs. CB control) x 4 types of tests (immediate text level, immediate inference level, delayed text level, and delayed inference level). The four immediate test comparisons were analyzed using two-level repeated measures analysis of variance. The four delayed tests were analyzed using two-level repeated measures analysis of covariance in order to control for the variance accounted for by performance on the immediate tests.

Table 4.1 shows a summary of the descriptive statistics, significance levels, and effect sizes for the overall comprehension comparisons.

TABLE 4.1

DESCRIPTIVE STATISTICS FOR COMPARISONS

*$p < .05$, $^{+}$ $p < .1$		MW vs. Self-Explanation Control				MW vs. Content Break Control			
		MW	SE Control			MW	CB Control		
Overall Comprehension	N	*M (SD)*	*M (SD)*	*D*	*P*	*M (SD)*	*M (SD)*	*d*	*p*
Immediate									
Text Level	27	.640 (.129)	.686 (.142)	-.346	.100	.640 (.129)	.609 (.178)	.198	.478
Inference Level	27	.442 (.140)	.491 (.171)	-.317	.127	.442 (.140)	.464 (.136)	-.161	.407
Delayed									
Text Level	13	.528 (.098)	.533 (.107)	-.050	.540	.513 (.124)	.497 (.134)	.119	.330
Inference Level	13	.405 (.193)	.344 (.071)	.422	.062*	.385 (.197)	.436 (.135)	-.304	.919
Non-intervention Concepts									
Immediate									
Text Level	27	.630 (.161)	.671 (.168)	-.244	.293	.630 (.161)	.600 (.194)	.171	.508
Inference Level	27	.439 (.151)	.474 (.214)	-.190	.356	.439 (.151)	.437 (.141)	.008	.971
Delayed									
Text Level	13	.536 (.225)	.562 (.254)	-.110	.157	.554 (.288)	.535 (.243)	.071	.955
Inference Level	13	.425 (.189)	.348 (.133)	.469	.012*	.373 (.192)	.436 (.173)	-.344	.286
Intervention Concepts									
Immediate									
Text Level	27	.652 (.231)	.742 (.209)	-.409	.091*	.652 (.231)	.602 (.271)	.199	.393
Inference Level	27	.459 (.325)	.542 (.270)	-.278	.187	.459 (.325)	.546 (.311)	-.272	.250
Delayed									
Text Level	13	.537 (.159)	.513 (.118)	.168	.307	.508 (.157)	.491 (.132)	.116	.317
Inference Level	13	.323 (.346)	.378 (.317)	-.166	.274	.346 (.333)	.364 (.237)	-.060	.146

4.3.1 MW versus SE Control

First, participants in the MW condition and the SE yoked control condition were compared across the four comprehension measures. The MW condition was predicted to outperform the SE control condition on the comprehension measures. Repeated-measures ANOVAs revealed there were no statistically significant differences in the two conditions on either of the immediate comprehension tests, text level: $F(1,26) = 2.91$, $p = .100$; inference level: $F(1,26) = 2.48$, $p = .127$. Although not statistically significant ($p = .100$), the difference in the immediate text level comprehension scores between the SE control and the MW condition was consistent with a small effect size ($d = -.346$, SE control > MW condition; see Table 4.1), which is the opposite of my prediction. The MW and SE control condition did not differ in their delayed text level comprehension, $F(1,8) = .409$, $p = .540$. However, in line with predictions, a repeated measures ANCOVA revealed a marginally significant main effect of condition for the delayed inference level comprehension test, such that participants in the MW condition scored higher than participants in the SE control, $F(1,8) = 4.68$, $p = .062$. In sum, MW participants did not significantly outperform the SE controls as predicted, with the exception of the delayed inference level test, in which there was marginally significant effect.

4.3.2 MW versus CB Control

Next, I compared the MW condition to the CB control condition. Participants in the MW condition were predicted to score higher on the comprehension measures than the CB control. However, repeated measures ANOVAs comparing MW and CB conditions on the immediate text and inference level comprehension measures revealed no significant effect of condition; text level $F(1,26) = .517$, $p = .478$; inference level $F(1,26) = .709$, $p = .407$. A repeated measures ANCOVA also revealed no main effect of condition for the delayed text level comprehension test, $F(1,8) = 1.08$, $p = .330$. Finally, the same null effect of condition was found for delayed inference level comprehension, $F(1,8) = .011$, $p = .919$.

Taken together, only one significant difference was found in overall comprehension across the eight tests performed: participants in the MW condition outperformed their SE yoked counterparts on the delayed inference level comprehension test.

4.4 Do Interventions Moderate the Effects of Condition?

The goal of research question 3 was to determine how the interventions influenced comprehension. Participants read a total of 15 concepts in the text, and each concept was linked to particular questions on the comprehension measures. The questions were split into two groups based on whether there was an intervention on the concept or not. Comprehension scores were then calculated separately for items that corresponded to intervention concepts versus non-intervention concepts. As in research question 2, a two level repeated measures analysis approach was used to test for condition differences in the four comprehension measures. The only difference was that separate models were computed for the intervention concepts and non-intervention concepts. Table 4.1 contains the descriptive statistics, effect sizes, and p-values for all comparisons.

4.4.1 MW versus SE Control

The MW condition was predicted to score significantly higher than the SE control on comprehension measures that corresponded to non-intervention concepts, but no comprehension differences were expected on intervention concepts.

Non-intervention concepts. Repeated measures ANOVAs revealed no effect of condition on comprehension for non-intervention concepts on immediate text level comprehension, $F(1,26) = 1.51$, $p = .293$, or immediate inference level comprehension, $F(1,26) = .356$, $p = .356$.

A repeated measures ANCOVA also yielded no effect of condition on delayed text level comprehension for non-intervention concepts, $F(1,8) = 2.44$, $p = .157$. There was, however, a significant main effect of condition for delayed inference level comprehension on non-intervention concepts, $F(1,8) = 10.59$, $p = .012$. Participants in the MW condition performed significantly better their SE control pairs, suggesting that the difference found for overall inference level comprehension might, in part, be attributed to the non-intervention concepts.

Intervention concepts. A repeated measures ANOVA revealed a marginally significant effect of condition on comprehension for intervention concepts, such that the SE control condition performed better than the MW condition, $F(1,26) = 3.08$, $p = .091$. There were no significant main effects of condition for immediate inference level comprehension, $F(1,26) = 1.83$, $p = .187$. Repeated measures ANCOVAs also showed no significant condition differences on delayed text level comprehension, $F(1,8) = 1.19$, $p = .307$, or delayed inference level comprehension for intervention concepts, $F(1,8) = 1.38$, $p = .274$.

To summarize, the SE control outperformed the MW condition on immediate text level comprehension (marginally significant) for intervention concepts, whereas the MW condition scored higher on the delayed inference level comprehension test for non-intervention concepts.

4.4.2 MW versus CB Control

The MW condition was expected to outperform the CB control condition on both intervention and non-intervention concepts. The same set of analyses (conducted for MW versus SE control) was repeated in order to test for a main effect of condition (MW vs. CB control) on comprehension on the intervention and non-intervention concepts. No main effects of condition were found on any of the comprehension measures, suggesting that comprehension between the MW and CB control conditions were similar regardless of the presence or absence of interventions, all p’s > .100.

CHAPTER 5:

DISCUSSION

5.1 Discussion

Much effort has been devoted to preventing mind wandering from occurring in the first place, but few interventions have been designed to reactively respond to mind wandering episodes as they arise to improve comprehension. As a step in this direction, a real-time intervention was developed and tested in order to combat the negative effects of mind wandering during reading. An experimental condition (MW condition) received automated self-explanation interventions when they were predicted to be mind wandering. This condition was compared to two different yoked control conditions (self-explanation and content-break control) that received interventions or reading breaks at the same points in the text. The main findings suggest that the MW condition had better inference level comprehension compared to the SE control condition when tested after a delay, and that this difference was limited to information on non-intervention concepts. Meanwhile, the MW condition and CB control did not significantly differ on any measure of comprehension. In this section, I address the extent to which my predictions were supported, interpret the results in the context of the cascade model of inattention, and address both limitations and future avenues of this research.

5.2 Alignment of Findings with Predictions

The cascade model of inattention suggests that mind wandering impedes comprehension at basic levels of processing. Therefore, a major goal of the self-explanation intervention was to repair breakdowns in processing at the conceptual level. The intervention was designed to accomplish this goal in two stages. First, it assessed participants' understanding of the main idea through automatically scoring their self-explanations. Second, participants were asked to re-read/re-answer if they did not demonstrate adequate understanding in step one. In line with predictions (see Figure 2.3), the self-explanation interventions improved self-explanation scores in the MW condition. However, the SE control condition did not outperform the control conditions after such interventions.

These data provide evidence that the mind-wandering detector was in fact sensitive to MW participants' attentional engagement: in effect, the detector predicted when comprehension could be improved by flagging periods of inattention to the text. Moreover, the increase in self-explanation scores specific to the mind-wandering targeted interventions also suggests, for the first time, that it is possible to detect and intervene to improve conceptual comprehension during the reading process. Indeed, repairing this type of basic online comprehension in the moment may be sufficient for some tasks, such as some forms of copy-editing.

Improving scores in real-time relied on the combined success of the mind wandering detector, the word overlap scoring algorithm, and the self-explanation questions. Therefore, this dissertation provides an outline for the types of challenges that go into designing a real-time mind wandering intervention, and offers some evidence that it is possible.

5.2.1 MW versus SE Control

Interventions repaired comprehension at the concept level, but how does that transfer to performance on the comprehension tests? The cascade model of inattention predicts that timely self-explanation interventions should improve overall comprehension in the MW condition. The interventions give participants a chance to fill in any gaps in their situation model as they read, effectively avoiding a cascading effect of impaired comprehension as readers progress through the text (Smallwood, 2011). The SE control condition (unlike the MW condition) did not experience mind wandering sensitive interventions, nor was their situation model repaired in the same way. The MW condition was therefore predicted to perform better than the SE control on the posttests after reading. Results did not support this prediction with one exception: MW performed (marginally significantly) better than the SE control condition on the delayed inference level test.

The difference between the MW and SE control conditions on the delayed inference level test aligns with Smallwood's (2011) prediction that mind wandering can influence comprehension at multiple levels. Here, repairing deficits in comprehension as they arise during reading helped to promote a long-term, deeper-level understanding of the text. While Smallwood did not necessarily predict that benefits would only emerge after (one week) delays, a consolidation period often supports achieving deeper level of conceptual understanding (Smallwood, 2011; Smallwood, Fishman, et al., 2007).

Splitting the 15 conceptual units into intervention concept and non-intervention concepts sheds additional light on how the interventions influenced comprehension in the MW and SE conditions differently over time. The SE control condition performed better than the MW condition on the immediate text level comprehension test that was associated with intervention concepts. This pattern can be attributed to the fact that self-explanation interventions in the SE control served to boost immediate information uptake on intervention concepts, as these participants were likely to be paying attention while reading about those concepts. Indeed, re-exposure may have simply strengthened their surface level comprehension when given the chance to re-read the text that they had already absorbed quite well (Rawson, Dunlosky, & Thiede, 2000). However, this benefit did not withstand the test of time: The MW and SE control were not significantly different in text level comprehension of intervention concepts one week later. The benefits to the SE control participants may not have been sustained because retention of lower-level textual information is often less durable due to low prioritization at encoding (Kintsch, Welsch, Schmalhofer, & Zimny, 1990).

An opposite pattern was found for the non-intervention concepts: participants in the MW condition outperformed the SE control on the delayed inference comprehension test. The non-intervention concepts highlight the concepts presented when participants in the MW condition were unlikely to have been mind wandering. The cascading model of inattention suggests that for these non-intervention concepts, the multiple levels of comprehension were not being interrupted or affected by lapses in attention. Since their mental model was repaired on concepts when they were mind wandering, their broader comprehension on other concepts was not strongly affected by their mind wandering episodes. Participants in the SE condition, on the other hand, may have experienced mind wandering on the non-intervention concepts. Their models of the text would have gone unrepaired, and they may ultimately have comprehended less about the concepts.

Taken together, the effect sizes (Cohen's ds) shown in Figure 5.1 suggest that there was a trend in which the SE condition scored higher on the immediate tests (though this effect was not significant), an in which the effects were reduced or even reversed when participants were tested after a delay. A takeaway from these comparisons is that the MW and SE control had different reading experiences due to how the detector was deployed (contingently and non-contingently with respect to participants' mind wandering). The SE control performed well on the immediate comprehension measures for intervention concepts, but did not exhibit any long term benefits from the extra exposure. By the one week delay, the MW condition was up to par on intervention-related comprehension and showed better performance on the delayed inference level tests, particularly for nonintervention concepts, after the consolidation period.

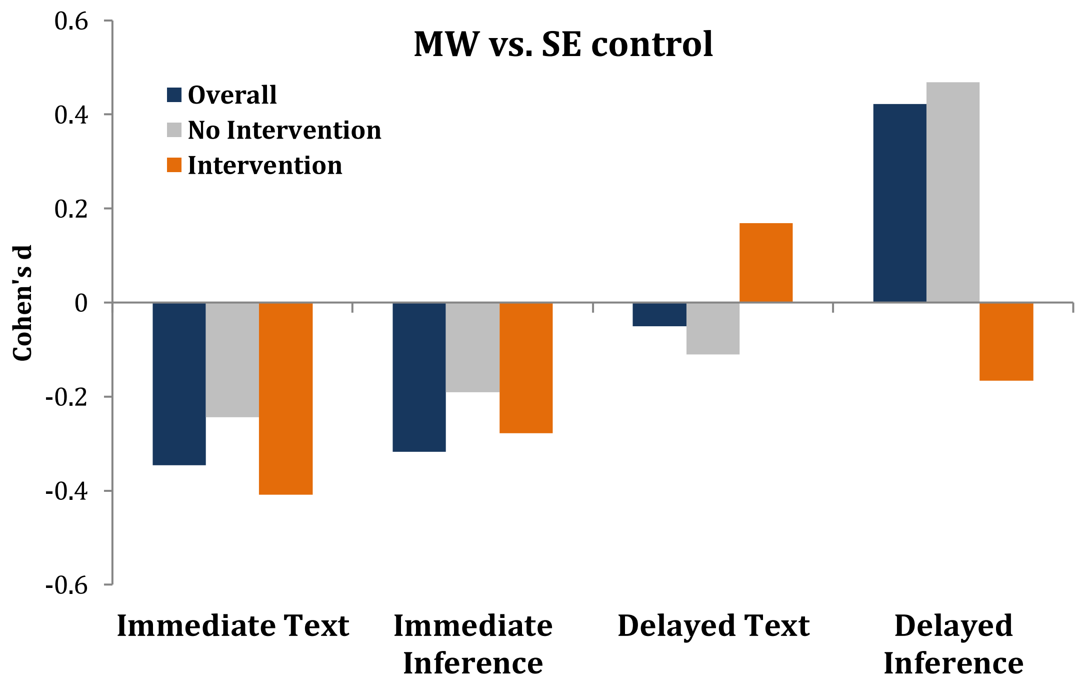


Figure 5.1: Graphical representation of Cohen's *d* for the MW vs. SE control condition comparisons for overall, intervention, and non-intervention concepts. Higher values correspond to MW having higher scores and lowers values mean control condition having high scores. Values closer to the x-axis mean similar values were found across conditions.

There are two lines of research that may help explain why the MW condition only performed better after a one week delay. First, research on retrieval induced forgetting (RIF) suggests that participants in SE control condition may have essentially forgotten related information once they successfully recalled it on the immediate tests. The RIF paradigm posits that memory for competing information can be suppressed, and that this effect can last for up to a week (Abel & Bäuml, 2012; Murayama, Miyatsu, Buchli, & Storm, 2014). Thus, when the SE control was tested on specific information on the immediate tests, competing information (e.g., for a given concept, the information that was not explicitly tested, such as other text level information) was suppressed at that point. As such, they performed worse on the delayed tests because the information was still suppressed. To confirm this idea in a follow-up study, participants might receive identical items on both the immediate and delayed tests instead of getting alternate forms with different items.

Second, the two conditions may have had different experiences with the immediate tests, subsequently leading to effects on the delayed test. This phenomenon is referred to the testing effect (Pavlik & Anderson, 2005; Whitten & Bjork, 1977). This position would suggest that the MW condition found the immediate tests relatively difficult, which lead to deeper processing of the information and perhaps better consolidation over time. Conversely, the SE control may have found the immediate tests to be fairly easy, as they were able to recall more text level information. Thus, they may not have processed the information encountered on the tests more deeply because they did not experience any increased difficulty, and ultimately this may have led to poorer encoding and thus a shorter lifespan (Bjork & Allen, 1970). This is also in line with the differences found for the non-intervention and intervention concepts because answering an immediate question in the intervention condition should be easier, as the participants re-read the information. In the non-intervention concept, the immediate testing might have positively affected delayed retention, as deeper levels of processing were involved with answering the immediate question when the text was not part of a self-explanation intervention (no re-exposure).

5.2.2 MW versus CB Control

Comparing the MW and CB control condition yielded a different pattern of results in that they did not differ statistically ($p > .100$) on any of the comprehension measures tested. This was also the case when comprehension was split by intervention and non-intervention pages. These findings did not align with the predicted patterns. The MW condition was expected to perform better since interventions for the CB condition were not mind wandering sensitive, nor did they get any self-explanation practice or re-exposure.

The lack of differences between the conditions is particularly compelling for the positive effects of periodic event breaks on comprehension. The event horizon model proposes that information is transferred from working memory to longer-term memory when an event boundary is encountered (Radvansky, 2012). In the present study, breaks always occurred at the end of a concept. It might be the case that any clear content break signifying the end of a conceptual unit helped the reader remember the information associated with that unit, clearing working memory before the onset of the next concept. Indeed, various types of “mental breaks” improve performance in other domains as well. For example, giving participants a break can alleviate vigilance decrement (Ariga & Lleras, 2011; Helton & Russell, 2015; Ralph, Onderwater, Thomson, & Smilek, 2016).

The CB condition performed especially well on the delayed tests. One explanation for this may be due to a lack of interference during the interventions in the CB condition. Specifically, interference occurs when the ability to remember an item is weakened due to its similarity to other items already stored memory (Anderson & Neely, 1996). The CB condition received a content break, where they were not re-exposed to text and they did not have to retrieve any text-related information from memory to generate a self-explanation. Therefore, their memories of the text may have been recalled more effectively after a consolidation period in the absence of interfering text-related information during the interventions. In particular, the interventions may have triggered selective retrieval of interfering information, which could harm later recall according to Anderson & Neely (1996).

5.3 Theoretical Implications

There is clear correlational evidence that mind wandering is negatively related to comprehension (Randall et al., 2014; Smallwood, Fishman, et al., 2007; Szpunar, Moulton, et al., 2013). The cascade model of inattention helps explain this phenomenon by proposing that "consequences of poor encoding could 'cascade' downward through the cognitive system, so that simple deficits in superficial processing could lead to more obvious deficits at a deeper level of analysis" (Smallwood, et al., 2007, p. 233). However, at this point, the cascade model has only been explicitly tested using word or sentence level reading paradigms and evaluating more shallow levels of comprehension (Smallwood, Fishman, et al., 2007; Smallwood, McSpadden, & Schooler, 2008). For example, if participants were mind wandering and missed a critical factual detail of the text, such as the clothing of the killer in a murder mystery, they would not be able to identify the killer later based on their outfit on a comprehension test.

The cascade model posits that if readers could make up for the deficits at the level of encoding, it could stop the cascading negative effects of mind wandering that impair comprehension. In this experiment, I attempted to do just that: repair deficits as they arise to promote deeper levels of comprehension using a self-explanation intervention. My findings suggest that influencing overall comprehension, or understanding of both intervention concepts and non-intervention concepts, is not as simple as filling in the gaps of missed conceptual information, as the cascade model might suggest (e.g., making sure they know what the killer was wearing). Benefits in the MW condition were limited to non-intervention concepts, suggesting that filling in the gaps of any missed information provides a benefit, but perhaps by enabling deeper processing on concepts when they are not mind wandering rather than comprehension of concepts encountered during mind wandering. It is possible that by helping MW participants build more complete mental models, the intervention assuaged the negative cascading effect from influencing comprehension other parts of the text.

Moreover, the comprehension benefits of a mind wandering sensitive intervention were only seen over a time delay, and the SE control condition still outperformed the MW condition on immediate text level comprehension for intervention concepts. This suggests that a gap in processing due to mind wandering may still have some immediate negative effects, even with receiving the self-explanation interventions. One idea is that the “breakdown in the top-down control of comprehension” (Smallwood et al., 2008, p. 1149) that coincides with mind wandering during reading is simply not easy to rebound from cognitively. A reader may adopt a new method of processing upon re-exposure to the material during the intervention. In this case, despite having the chance to re-read and self-explain, they may have lacked top-down motivation to process the lower levels of the text, and instead just looking for the main ideas may have been sufficient. For this reason, it is important to test other intervention methods. For example, the intervention here targeted deeper processing through self-explanation which required participants to get the overall gist of a concept. Other intervention methods might include asking participants come up with their own questions about a concept or asking them to answer application questions that would require transfer knowledge.

More work is needed to better understand the precise negative cascading effects of mind wandering on comprehension, especially on inference level comprehension. This project, however, provides some evidence that preventing the cascading effect might positively influence comprehension over time, and provides some insights into the mechanisms by which such an intervention influences such comprehension processes.

5.4 Limitations and Future Directions

It is important to note the limitations of the current experiment. First, power was an issue for many of the comparisons in Table 4.1. This was particularly true for comparisons involving the delayed tests. Future research could include replication with a larger sample, as well as using other methods of doing the delayed testing to encourage (e.g., online surveys).

Second, only one text was used in this study. This text was written a century ago and it was likely unfamiliar to any of our participants. However, it is possible that it elicited specific reading behaviors that would not generalize to more current texts, or to texts that were extremely interesting to participants. Future work should evaluate mind wandering interventions using more varied texts, and across different learning contexts.

For example, participants read for less than thirty minutes in the current study in a single session. One follow up question would be how does a similar detector work over time, across multiple sessions? This would be particularly interesting when comparing the MW condition and SE control over a long narrative text that requires a very strong situational model in order to understand the text over time. It would also be an interesting comparison for the MW condition and CB control. Do these types of breaks in reading consistently help over a long period of time? Since mind wandering interventions rely on successful detectors, this idea and other future studies will also rely on new (or adapted) mind wandering detectors. The field of mind wandering detection is rapidly evolving, so building new detectors and testing more interventions is a promising avenue of research.

A third limitation is the cost of the eye tracking technology used in this experiment, making scalability an issue outside the laboratory. Fortunately, cost-effective eye tracking are increasing available (e.g., Eye Tribe and Tobii EyeX). As reliable mind wandering detectors are built using more scalable technologies, follow-up studies can be implemented using more affordable eye-tracking methods. Finally, this was a lab study with a sample of 81 students from one university. The findings from this study should be interpreted and applied with caution, as it was conducted in a laboratory setting this setting is not representative of an ecological classroom, and with participants that were homogenous in their age and other demographic characteristics relative to the general population for which such interventions might be relevant.

5.5 Concluding Remarks

The goal of my dissertation was to test an automatic intervention designed to combat the cascading negative effects of mind wandering on reading comprehension. Two major goals were accomplished by conducting this research: (1) I developed a self-explanation intervention that aimed to promote deep processing of concepts presented in a written text, and (2) I tested the intervention in a yoked-control experiment. Results indicate that the real-time intervention helped boost comprehension during reading when mind wandering was detected. Furthermore, the results suggest that mind wandering-sensitive self-explanation interventions indeed promoted better long-term conceptual comprehension compared to pseudorandom self-explanation interventions, particularly for concepts when there was no intervention present. However, contrary to my predictions, participants who received mind wandering-sensitive interventions did not score higher on the comprehension tests compared to the participants who simply received periodic breaks from reading.

In sum, these data suggest that it is to implement interventions during learning by tracking real-time behaviors that indicate mind wandering. Furthermore, such an intervention may help optimize student's learning experience when they inevitably start to mind wander. This dissertation serves as just one of the initial explorations on this front. Promising next steps in this line of research might include explorations of different learning/intervention strategies coupled with improved detection methods, and other experimental designs for evaluating the efficacy of real-time interventions for mind wandering. In addition, future work on this topic can help further understand and refine the mechanisms influencing comprehension suggested by the cascade model of inattention.

APPENDIX A:

EXPERIMENTAL TEXT

As the earlier editions of this book have met with so favorable a reception, since in fact about two tons of my bubbles are floating about the world, and the book has been translated into French, German and Polish, I have thought fit to rearrange, alter and enlarge it. The chapter on the colors and thicknesses of bubbles is entirely new, as are two or three other shorter ones on bubbles of different kinds. In some of these, especially that on their colors, the treatment of the subject is necessarily a good deal more difficult than it is in the original parts. As the book is primarily intended for the general reader rather than for the student of physical science I have avoided the use of all trigonometrical and algebraical formulae, as I know their paralyzing effect on the non-technical reader. At the same time I do not think that there is any want of precision or accuracy as a result. I have therefore been compelled to employ a more cumbersome arithmetical treatment in some cases, while in others I have used geometrical construction in order to obtain quantitative results. This has the advantage of providing ocular demonstration as well as proof, and in the case of the loss of time within the thickness of a soap film is far neater and more natural than the more usual trigonometrical method. I have felt constrained to use the archaic British units of measurement, as the unfamiliar metric terminology would have distracted the attention of the majority for whom this book is intended, who have spent untold hours that might have gone into mathematical or general education in performing ridiculous operations such as reduction, compound multiplication and practice which our British methods of measurement necessitate, but which in more enlightened countries are wholly unnecessary. This book is not prepared to meet the requirements and artificial restrictions of any syllabus, and it is not prepared to help students through any examination. I cannot help thinking, however, that if the type of student who puts more faith in learning formulae than in understanding how they may be recovered when forgotten, as they will be, would condescend to spend the time necessary for reading the chapter on the colors of soap-bubbles he would derive some help from it, and he might even find it useful in preparing for an examination. In the additional chapters I have found it more convenient to give with the text sufficient guidance for the repetition of experiments instead of reserving this for the practical hints at the end, which remain much as they were. I do not suppose that there is anyone in this room who has not occasionally blown a common soap-bubble, and while admiring the perfection of its form, and the marvelous brilliancy of its colors, wondered how it is that such a magnificent object can be so easily produced. I hope that none of you are yet tired of playing with bubbles, because, as I hope we shall see, there is more in a common bubble than those who have only played with them generally imagine. The wonder and admiration so beautifully portrayed by Millais in a picture, copies of which, thanks to modern advertising enterprise, some of you may possibly have seen, will, I hope, in no way fall away in consequence of these lectures; I think you will find that it will grow as your knowledge of the subject increases. Plateau in his famous work, Statique des Liquides, quotes a passage from a book by Henry Berthoud, to the effect that there is an Etruscan vase in the Louvre in Paris in which children are represented blowing bubbles from a pipe. Plateau states, however, that no classical author refers to any such amusement, and the only two references to bubbles of any kind that he can find are in Ovid and Martial. I have hunted for this vase at the Louvre in vain. A correspondent, however, sent the quotation to the director, by whom he was informed that no such vase was there, but that a number of fictitious antique vases had been removed from the collection. It is possible that some of you may like to know why I have chosen soap-bubbles as my subject; if so, I am glad to tell you. Though there are many subjects which might seem to a beginner to be more wonderful, more brilliant, or more exciting, there are few which so directly bear upon the things which we see every day. You cannot pour water from a jug or tea from a tea-pot; you cannot even do anything with a liquid of any kind, without setting in action the forces to which I am about to direct your attention. You cannot then fail frequently to be reminded of what you will hear and see in this room, and, what is perhaps most important of all, many of the things I am going to show you are so simple that you will be able without any apparatus to repeat for yourselves the experiments which I have prepared, and this you will find more interesting and instructive than merely listening to me and watching what I do.

There is one more thing I should like to explain, and that is why I am going to show experiments at all. You will at once answer, because it would be so dreadfully dull if I didn't. Perhaps it would. But that is not the only reason. I would remind you then that when we want to find out anything that we do not know, there are two ways of proceeding. We may either ask somebody else who does know, or read what the most learned men have written about it, which is a very good plan if anybody happens to be able to answer our question; or else we may adopt the other plan, and by arranging an experiment, find out for ourselves. An experiment is a question which we ask of Nature, who is always ready to give a correct answer, provided we ask properly, that is, provided we arrange a proper experiment. An experiment is not a conjuring trick, something simply to make you wonder, nor is it simply shown because it is beautiful, or because it serves to relieve the monotony of a lecture; if any of the experiments I show are beautiful, or do serve to make these lectures a little less dull, so much the better; but their chief object is to enable you to see for yourselves what the true answers are to the questions that I shall ask. Now I shall begin by performing an experiment which you have all probably tried dozens of times without recognizing that you were making an experiment at all. I have in my hand a common camel's-hair brush. If you want to make the hairs cling together and come to a point, you wet it, and then you say the hairs cling together because the brush is wet. Now let us try the experiment; but as you cannot see this brush across the room, I hold it in the lantern, and you can see it enlarged upon the screen. Now it is dry, and the hairs are separately visible. I am now dipping it in the water, as you can see, and on taking it out, the hairs, as we expected, cling together, because they are wet, as we are in the habit of saying. I shall now hold the brush in the water, but there it is evident that the hairs do not cling at all, and yet they surely are wet now, being actually in the water. It would appear then that the reason which we always give is not exactly correct. This experiment, which requires nothing more than a brush and a glass of water, then, shows that the hairs of a brush cling together not only because they are wet, but for some other reason as well which we do not yet know. It also shows that a very common belief as to opening our eyes under water is not founded on fact. It is very commonly said that if you dive into the water with your eyes shut you cannot see properly when you open them under water, because the water gums the eyelashes down over the eyes; and therefore you must dive in with your eyes open if you wish to see under water. Now as a matter of fact this is not the case at all; it makes no difference whether your eyes are open or not when you dive in, you can open them and see just as well either way. In the case of the brush we have seen that water does not cause the hairs to cling together or to anything else when under the water, it is only when taken out that this is the case. This experiment, though it has not explained why the hairs cling together, has at any rate told us that the reason always given is not sufficient. I shall now try another experiment as simple as the last. I have a pipe from which water is very slowly issuing, but it does not fall away continuously; a drop forms which slowly grows until it has attained a certain definite size, and then it suddenly falls away. I want you to notice that every time this happens the drop is always exactly the same size and shape. Now this cannot be mere chance; there must be some reason for the definite size and shape. Why does the water remain at all ? It is heavy and is ready to fall, but it does not fall; it remains clinging until it is a certain size, and then it suddenly breaks away, as if whatever held it was not strong enough to carry a greater weight. Mr. Worthington has carefully drawn on a magnified scale the exact shape of a drop of water of different sizes, and these you now see upon the diagram on the wall. These diagrams will probably suggest the idea that the water is hanging suspended in an elastic bag, and that the bag breaks or is torn away when there is too great a weight for it to carry. It is true there is no bag at all really, but yet the drops take a shape which suggests an elastic bag. To show you that this is no fancy, I have supported by a tripod a large ring of wood over which a thin sheet of india-rubber has been stretched, and now on allowing water to pour in from this pipe you will see the rubber slowly stretching under the increasing weight, and, what I especially want you to notice, it always assumes a form like those on the diagram. As the weight of water increases the bag stretches, and now that there is about a pail full of water in it, it is getting to a state which indicates that it cannot last much longer; it is like the water-drop just before it falls away, and now suddenly it changes its shape, and it would immediately tear itself away if it were not for the fact that india-rubber does not stretch indefinitely; after a time it gets tight and will withstand a greater pull without giving way. You therefore see the great drop now permanently hanging which is almost exactly the same in shape as the water drop at the point of rupture. I shall now let the water run out by means of a siphon, and then the drop slowly contracts again. Now in this case we clearly have a heavy liquid in an elastic bag, whereas in the drop of water we have the same liquid but no bag that is visible. As the two behave in almost exactly the same way, we should naturally be led to expect that their form and movements are due to the same cause, and that the small water-drop has something holding it together like the india-rubber you now see. Let us see how this fits the first experiment with the brush. That showed that the hairs do not cling together simply because they are wet; it is necessary also that the brush should be taken out of the water, or in other words it is necessary that the surface or the skin of the water should be present to bind the hairs together. If then we suppose that the surface of water is like an elastic skin, then both the experiments with the wet brush and with the water-drop will be explained. Let us therefore try another experiment to see whether in other ways water behaves as if it had an elastic skin. I have here a plain wire frame fixed to a stem with a weight at the bottom, and a hollow glass globe fastened to it with sealing-wax. The globe is large enough to make the whole thing float in water with the frame up in the air. I can of course press it down so that the frame touches the water. To make the movement of the frame more evident there is fixed it to a paper flag. Now if water behaves as if the surface were an elastic skin, then it should resist the upward passage of the frame which I am now holding below the surface. I let go, and instead of bobbing up as it would do if there were no such action, it remains tethered down by this skin of the water. If I disturb the water so as to let the frame out at one corner, then, as you see, it dances up immediately. You can see that the skin of the water must have been fairly strong, because a weight of about one quarter of an ounce placed upon the frame is only just sufficient to make the whole thing sink. This apparatus, which was originally described by Van der Mensbrugghe, I shall make use of again in a few minutes. I can show you in a more striking way that there is this elastic layer or skin on pure clean water. I have a small sieve made of wire gauze sufficiently coarse to allow a common pin to be put through any of the holes. There are moreover about eleven thousand of these holes in the bottom of the sieve. Now, as you know, clean wire is wetted by water, that is, if it is dipped in water it comes out wet; on the other hand, some materials, such as paraffin wax, of which paraffin candles are made, are not wetted or really touched by water, as you may see for yourselves if you will only dip a paraffin candle into water. I have melted a quantity of paraffin in a dish and dipped this gauze into the melted paraffin so as to coat the wire all over with it, but I have shaken it well while hot to knock the paraffin out of the holes. You can now see on the screen that the holes, all except one or two, are open, and that a common pin can be passed through readily enough. This then is the apparatus. Now if water has an elastic skin which it requires force to stretch, it ought not to run through these holes very readily; it ought not to be able to get through at all unless forced, because at each hole the skin would have to be stretched to allow the water to get to the other side. This you understand is only true if the water does not wet or really touch the wire. Now to prevent the water that I am going to pour in from striking the bottom with so much force as to drive it through, I have laid a small piece of paper in the sieve, and am pouring the water on to the paper, which breaks the fall. I have now poured in about half a tumbler of water, and I might put in more, I take away the paper but not a drop runs through. If I give the sieve a jolt then the water is driven to the other side, and in a moment it has all escaped. Perhaps this will remind you of one of the exploits of our old friend Simple Simon, "who went for water in a sieve, but soon it all ran through." But you see if you only manage the sieve properly, this is not quite so absurd as people generally suppose. If now I shake the water off the sieve, I can, for the same reason, set it to float on water, because its weight is not sufficient to stretch the skin of the water through all the holes. The water, therefore, remains on the other side, and it floats even though, as I have already said, there are eleven thousand holes in the bottom, any one of which is large enough to allow an ordinary pin to pass through. This experiment also illustrates how difficult it is to write real and perfect nonsense. You may remember one of the stories in Lear's book of Nonsense Songs. They went to sea in a sieve, they did, In a sieve they went to sea: In spite of all their friends could say, On a winter's morn, on a stormy day, In a sieve they went to sea... They sailed away in a sieve, they did, In a sieve they sailed so fast, With only a beautiful pea-green veil, Tied with a rib and by way of a sail, To a small tobacco-pipe mast;" And so on. You see that it is quite possible to go to sea in a sieve--that is, if the sieve is large enough and the water is not too rough and that the above lines are now realized in every particular. I may give one more example of the power of this elastic skin of water. If you wish to pour water from a tumbler into a narrow-necked bottle, you know how if you pour slowly it nearly all runs down the side of the glass and gets spilled about, whereas if you pour quickly there is no room for the great quantity of water to pass into the bottle all at once, and so it gets spilled again. But if you take a piece of stick or a glass rod, and hold it against the edge of the tumbler, then the water runs down the rod and into the bottle, and none is lost; you may even hold the rod inclined to one side, as I am now doing, but the water runs down the wet rod because this elastic skin forms a kind of tube which prevents the water from escaping. This action is often made use of in the country to carry the water from the gutters under the roof into a water-butt below. A piece of stick does nearly as well as an iron pipe, and it does not cost anything like so much. I think then that I have now done enough to show that on the surface of water there is a kind of elastic skin. I do not mean that there is anything that is not water on the surface, but that, the water while there acts in a different way to what it does inside, and that it acts as if it were an elastic skin made of something like very thin india-rubber, only that it is perfectly and absolutely elastic, which india-rubber is not. You will now be in a position to understand how it is that in narrow tubes water does not find its own level, but behaves in an unexpected manner. I have placed in front of the lantern a dish of water colored blue so that you may the more easily see it. I shall now dip into the water a very narrow glass pipe, and immediately the water rushes up and stands about half an inch above the general level. The tube inside is wet. The elastic skin of the water is therefore attached to the tube, and goes on pulling up the water until the weight of the water raised above the general level is equal to the force exerted by the skin. If I take a tube about twice as big, then this pulling action which is going on all round the tube will cause it to lift twice the weight of water, but this will not make the water rise twice as high, because the larger tube holds so much more water for a given length than the smaller tube. It will not even pull it up as high as it did in the case of the smaller tube, because if it were pulled up as high the weight of the water raised would in that case be four times as great, and not only twice as great, as you might at first think. It will therefore only raise the water in the larger tube to half the height, and now that the two tubes are side by side you see the water in the smaller tube standing twice as high as it does in the larger tube. In the same way, if I were to take a tube as fine as a hair the water would go up ever so much higher. It is for this reason that this is called Capillarity, from the Latin word capillus, a hair, because the action is so marked in a tube the size of a hair. Supposing now you had a great number of tubes of all sizes, and placed them in a row with the smallest on one side and all the others in the order of their sizes, then it is evident that the water would rise highest in the smallest tube and less and less high in each tube in the row, until when you came to a very large tube you would not be able to see that the water was raised at all. You can very easily obtain the same kind of effect by simply taking two squares pieces of window glass and placing them face to face with a common match or small fragment of anything to keep them a small distance apart along one edge while they meet together along the opposite edge. An india-rubber ring stretched over them will hold them in this position. I now take such a pair of plates and stand it in a dish of colored water, and you at once see that the water creeps up to the top of the plates on the edge where they meet, and as the distance between the plates gradually increases so the height to which the water rises gradually gets less, and the result is that the surface of the liquid forms a beautifully regular curve which is called by mathematicians a rectangular hyperbola. I shall have presently to say more about this and some other curves, and so I shall not do more now than state that the hyperbola is formed because as the width between the plates gets greater the height gets less, or, what comes to the same thing, because the weight of liquid supported at any small part of the curve is always the same. If the plates or the tubes had been made of material not wetted by water, then the effect of the tension of the surface would be to drag the liquid away from the narrow spaces, and the more so as the spaces were narrower. As it is not easy to show this well with paraffined glass plates or tubes and water, I shall use another liquid which does not wet or touch clean glass, namely, quicksilver. As it is not possible to see through quicksilver, it will not do to put a narrow tube into this liquid to show that the level is lower in the tube than in the surrounding vessel, but the same result may be obtained by having a wide and a narrow tube joined together. Then, as you see upon the screen, the quicksilver is lower in the narrow than in the wide tube, whereas in a similar apparatus the reverse is the case with water. Although the elastic tension which I have called the strength of the water-skin is very small where big things are considered, this is not the case where very small things are acted upon. For instance, those of you who are fortunate enough to live in the country and who have gone down to play by the side of a brook must often have seen water spiders and other small creatures running on the water without sinking in. For some reason their feet are not wetted by the water, and so they tread down and form a small dimple where each foot rests, and so the up-pulling sides of the dimples just support the weight of the creature. It follows also from this that this weight is exactly equal to the weight of the water that would just fill all the dimples up to the general water level, that is, supposing that it were possible to imagine the dimpled water solidified for the purpose of the experiment. Mr. H. H. Dixon, of Dublin, once very ingeniously measured the force with which one of these water-spiders pressed its different feet while running on the water. He photographed the shadow of the spider and of the dimples upon a white porcelain dish containing the water on which the spider ran. He then mounted one of the spider's feet on a very delicate balance, and made it press on the water with different degrees of pressure, and again photographed the shadow of the dimple for each degree of pressure. He was thus able to make a scale by means of which he could tell the pressure exerted by the spider on any foot by comparing the size of the shadow of the dimples with those on his scale. He was also able to see the order in which the water-spider put down its feet, and so solve the problem for the spider which so perplexed the centipede of the well-known lines - "A centipede was happy quite, Until a toad, in fun, Said, 'Pray which leg goes after which?' This raised her doubts to such a pitch She fell distracted in the ditch, Not knowing how to run." Professor Miall has described how a certain other water spider spins a net under water through which air will not pass, just as water would not pass through the sieve which it did not wet. The spider then goes to the surface and carries air down and liberates it under the net and so gradually accumulates a reservoir of air to enable it to breathe at leisure when it has a good supply. The elastic water-skin is made use of by certain larvae which live immersed in the water, as well as by water spiders and other creatures which run over it. The common gnat lays its eggs in stagnant water, and seems especially fond of water-butts and troughs in gardens or greenhouses. These eggs in time hatch out into the larva of the gnat which corresponds to the silkworm or caterpillar of a moth or butterfly. This larva may be found in thousands in hot weather in ordinary rainwater- butts. You must have noticed small, dark creatures which swim in a curious jerky way and go to the bottom when you suddenly come near and frighten them. However, if you keep still, you will not have long to wait before you see them swimming back again to the surface, to which they attach themselves, and then remain hanging. It is very easy to show these alive on the screen. In the place of a lantern slide I have a cell containing water in which are a number of these larva, you will see how they swim to the surface and then hang by a projection like tail. This is a breathing tube, and so, even though they are heavier than water and naturally sink, they are able to hang by their breathing tubes, and breathe while they can eat rotten leaves without exertion. If you examine the surface of water, say in a tumbler in which you have placed some of these larva, you will see, when they come to the surface, as they will have to do, that there is a small dimple where they hang, and the weight of the water that would be needed to fill this dimple is exactly equal to the downward pull of the larva. If you look through a magnifying-glass you will see it much better and also notice what an interesting looking creature the larva really is. Professor Miall has also shown us how the common duckweed makes use of surface tension to turn itself round so that, if the pond is not already crammed with as many as it will hold, the leaves attract each other so as to rest in contact end to end, leaving their sides, from which the young plants originate by a process of budding, free for their development. The leaf has a central ridge, and the heel and toe, so to speak, are above the general surface, so that the water surface is curved up to these. Now it is a fact that things that are wetted by water so that the water is curved upwards where it meets them, attract one another, the reason being that the pressure within the raised water is less than the atmospheric pressure on the other side in proportion as it is raised higher. Also things neither of which is wetted by water attract each other, as in that case the water is curved downwards where it meets them, and the air pressure in the region of the depression is less than that of the water on the other side in proportion as it is depressed lower. On the other hand, two things, one of which is wetted and one of which is not wetted, repel each other. Now, coming back to the duckweed, the elevated toe and heel attract each other as wetted things do, while the sides, being depressed to the water level about, are neutral. You will do well to make the experiment yourself. Duckweed is easily found in the country. Fill a tumbler until it is rather over full, so that the curvature of the water at the edge is downwards. Small things that are wetted will then remain in the middle. Place a few duckweed plants in the water and see how they attract each other endways only, and how a wetted point will cause isolated ones to turn round almost like a magnet acting on a compass needle. Having got your duckweed, it will be found interesting to keep it a few days and watch the budding and separation of new plants. So far I have given you no idea what force is exerted by this elastic skin of water. Measurements made with narrow tubes, with drops, and in other ways, all show that it is almost exactly equal to the weight of three and a quarter grains to the inch. We have, moreover, not yet seen whether other liquids act in the same way, and if so whether in other cases the strength of the elastic skin is the same. You now see a second tube identical with that from which drops of water were formed, but in this case the liquid is alcohol. Now that drops are forming, you see at once that while alcohol makes drops which have a definite size and shape when they fall away, the alcohol drops are not by any means as large as the drops of water which are falling by their side. Two possible reasons might be given to explain this. Either alcohol is a heavier liquid than water, which would account for the smaller drop if the skin in each liquid had the same strength, or else if alcohol is not heavier than water its skin must be weaker than the skin of water. As a matter of fact alcohol is a lighter liquid than water, and so still more must the skin of alcohol be weaker than that of water. We can easily put this to the test of experiment. In the game that is called the tug-of-war you know well enough which side is the strongest; it is the side which pulls the other over the line. Let us then make alcohol, and water play the same game. In order that you may see the water, it is colored blue. It is lying as a shallow layer on the bottom of this white dish. At the present time the skin of the water is pulling equally in all directions, and so nothing happens; but if I pour a few drops of alcohol into the middle, then at the line which separates the alcohol from the water we have alcohol on one side pulling in, while we have water on the other side pulling out, and you see the result. The water is victorious; it rushes away in all directions, carrying a quantity of the alcohol away with it, and leaves the bottom of the dish dry. This difference in the strength of the skin of alcohol and of water, or of water containing much or little alcohol, gives rise to a curious motion which you may see on the side of a wine-glass in which there is some fairly strong wine, such as port. The liquid is observed to climb up the sides of the glass, then to gather into drops, and to run down again, and this goes on for a long time. This was explained by Professor James Thomson as follows: The thin layer of wine on the side of the glass being exposed to the air, loses its alcohol by evaporation more quickly than the wine in the glass does. It therefore becomes weaker. In alcohol or stronger in water than that below, and for this reason it has a stronger skin. It therefore pulls up more wine from below, and this goes on until there is so much that drops form, and it runs back again into the glass, as you now see upon the screen. It is probable that this movement is referred to in Proverbs xxiii. 31: "Look not thou upon the wine when it is red, when it gives his color in the cup, when it moved itself a right." Ether, in the same way, has a skin which is weaker than the skin of water. The very smallest quantity of ether on the surface of water will produce a perceptible effect. For instance, the wire frame which I left some time ago is still resting against the water-skin. The buoyancy of the glass bulb is trying to push it through, but the upward force is just not sufficient. I will however pour a few drops of ether into a glass, and simply pour the vapor upon the surface of the water (not a drop of liquid is passing over), and almost immediately sufficient ether has condensed upon the water to reduce the strength of the skin to such an extent that the frame jumps up out of the water. There is a well-known case in which the difference between the strength of the skins of two liquids may be either a source of vexation or, if we know how to make use of it, an advantage. If you spill grease on your coat you can take it out very well with benzene. Now if you apply benzene to the grease, and then apply fresh benzene to that already there, you have this result and there is then greasy benzene on the coat to which you apply fresh benzene. It so happens that greasy benzene has position a stronger skin than pure benzene. The greasy benzene therefore plays at tug-of-war with pure benzene, and being stronger wins and runs away in all directions, and the more you apply benzene the more the greasy benzene runs away carrying the grease with it. But if you follow the proper method, and first make a ring of clean benzene round the grease-spot, and then apply benzene to the grease, you then have the greasy benzene running- away from the pure benzene ring and heaping itself together in the middle, and escaping into the fresh rag that you apply, so that the grease is all of it removed. I put this to a very severe test once when a new white satin dress had been spoiled by an upset of soup. I laid a number of dusters on the ground out of doors and carefully laid the part which was stained over the dusters, and then using a quart or so of pure benzene (not benzoline or petrol which by comparison are useless). I poured it freely first in a ring round, and then through the place constantly, replacing the old dusters by new. Then on lighting it up the remaining benzene quickly evaporated, and no trace of the stain nor any "high-water mark" could be detected. There is a difference again between hot and cold grease, as you may see, when you get home, if you watch a common candle burning. Close to the flame the grease is hotter than it is near the outside. It has therefore a weaker skin, and so a perpetual circulation is kept up, and the grease runs out on the surface and back again below, carrying little specks of dust which make this movement visible, and making the candle burn regularly. You probably know how to take out grease-stains with a hot poker and blotting-paper. Here again the same kind of action is going on. A piece of lighted camphor floating on water is another example of movement set up by differences in the strength of the skin of water owing to the action of the camphor. The best way to make the experiment with camphor is to take a large basin of very clean water and then holding the camphor over the water, scrape the corners lightly with a knife. The minute specks of camphor which fall on the water will then display an activity which is quite surprising. The water however must be nearly perfectly free from grease, mere contact with the finger for a moment may be sufficient to stop the display of activity. Lord Rayleigh has determined the weight of oil that will "kill" camphor in water in a large bath, and has found in this way that if the thickness of oil is no more than tiny fraction of an inch, the camphor is dead.

APPENDIX B:

EXAMPLE READING SCREEN

As the earlier editions of this book have met with so favorable a reception, since in fact about two tons of my bubbles are floating about the world, and the book has been translated into French, German and Polish, I have thought fit to rearrange, alter and enlarge it. The chapter on the colors and thicknesses of bubbles is entirely new, as are two or three other shorter ones on bubbles of different kinds. In some of these, especially that on their colors, the treatment of the subject is necessarily a good deal more difficult than it is in the original parts. As the book is primarily intended for

APPENDIX C:

TEXT BROKEN DOWN BY CONCEPT

Section #	**Screens**	**Target Concept**
1	1-3	Conveying reason for unconventional methods to present quantitative results
2	4-7	Why bubbles are a historically good subject to focus on throughout the book
3	8-11	Description and examples of why experiments are important for understanding real-life phenomenon
4	12-14	What happens underwater with a brush/How experiments can show what is NOT a good explanation for a phenomenon
5	15-19	Elastic skin of water/What is the shape of a water droplet?
6	20-22	Elastic skin continued: Water can hold down items with impressive weight under water
7	23-25	Even with holes, something can float on water because of the elastic skin
8	26-28	Going to sea in a sieve is not unreasonable when certain conditions are met
9	29-32	How water behaves differently in different sized tubes (narrow versus wide)
10	33-36	Different liquids also behave differently, like when they do not actually wet the surface
11	37-40	How insects take advantage of surface tension to walk on water
12	41-45	How duckweed uses surface tension to attract/repel and reproduce
13	46-49	Tug of war game between water and alcohol: water wins because it is stronger
14	50-52	Wine stays on the sides of a glass because alcohol evaporates faster there
15	53-57	Pure benzene can get out grease stains through attracting the grease

APPENDIX D:

EXAMPLE STEP FUNCTION FOR DEPLOYING DETECTOR

Values that fall on the left side will not receive intervention (0 to .3)
Example: .20 = 0% chance of intervention

Values that fall in the middle will receive intervention probabilistically based on their average mind wandering likelihood. (.31 to .65)
Example: .40 = 40% chance of intervention

Values that fall on the right side will always receive intervention (.66 to 1.0)
Example: .77 = 100% chance of intervention

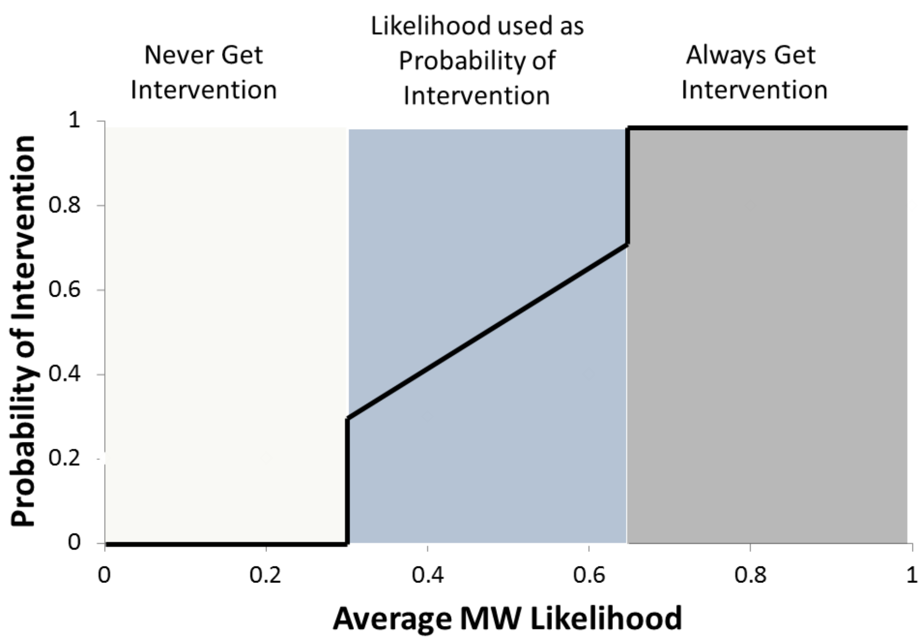


Figure D.1: Graphical representation of step function

APPENDIX E:

INTERVENTIONS QUESTIONS

1. Why did the author choose a different way to present the quantitative explanations in this book?
2. Why do you think the author focuses on bubbles for the topic of this book?
3. Why does the author describe experiments throughout the text?
4. Based on the hairbrush experiment, what would happen to your clothes under water if you jumped in a pool with them on?
5. How does the india-rubber experiment behave similar to water droplets?
6. How did the experiment with the globe demonstrate that water is like an elastic skin?
7. What point does the author make about when water will run through a sieve?
8. How does "going to sea in a sieve" seem possible based on the authors demonstration?
9. Why does the water rise higher in the in the small tubes?
10. What are the key differences in the results from water versus quicksilver demonstration?

11. Explain some of the ways the spider described in the text differs from other spiders?
12. Explain how the author relates surface tension to duckweed.
13. Explain what the author means by saying "the water is victorious"?
14. Describe the process the author thinks refers to wine's ability to move "itself a right"?
15. Explain how you should clean a grease stain with benzene?

APPENDIX F:

EXAMPLE INTERVENTION SCREENS

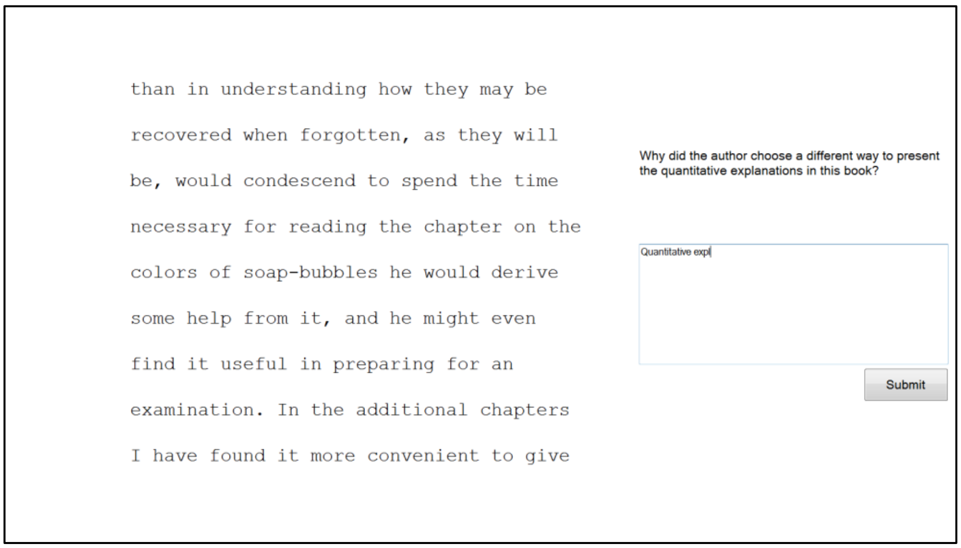


Figure F.1: Example Screen Shot of Intervention

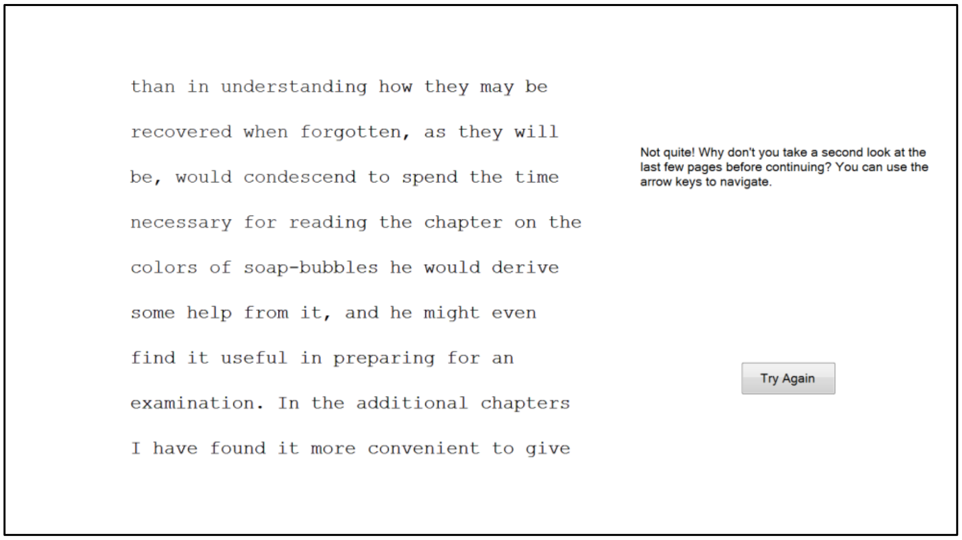


Figure F.2: Example Screen Shot of Intervention Feedback

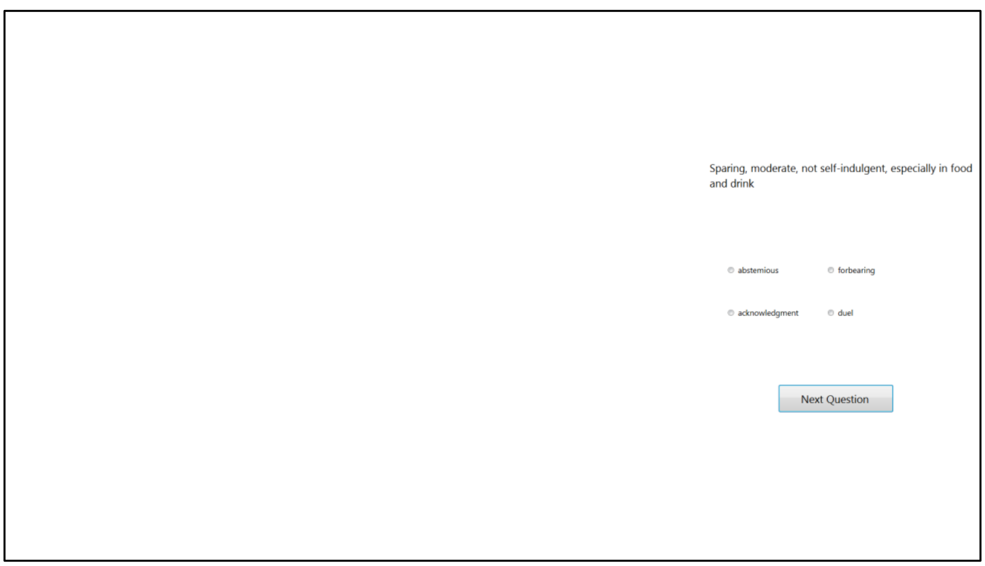


Figure F.3: Example Screen Shot of Tip-of-the-Tongue Questions

APPENDIX G:

TIP OF THE TONGUE QUESTIONS FROM MEYER & BOCK (1992)

Target Definition	Target	Semantic Clue	Phonological Clue	Unrelated Cue
To keep eggs warm until hatching	incubate	nurture	ignorant	floor
Of, on, or with two sides	bilateral	double	bewilderment	youth
Hollow or, as regards people, unintelligent, expressionless	vacuous	blank	victory	chamber
Sparing, moderate, not self-indulgent, especially in food and drink	abstemious	forbearing	acknowledgment	duel
Long close garment worn particularly by clergy and choir singers [choristers], often under a surplice	cassock	robe	careless	division
Something out of keeping with the time in which it exists	anachronism	premature	accelerator	view
To make calm or serene or to reduce agitation, especially by use of a drug	tranquilize	soothe	tropical	deep
Adherent of [to] the view that human action is not free but directed by external forces acting on the will	determinist	fatalist	diminutive	hall
Able to read and write	literate	intellectual	lavender	mind
The desire to do evil, or for others to have ill fortune	malevolence	hostile	molecular	drink
The study of the developing fetus before birth	embryology	cellular	electricity	night
Feeding on both plants and flesh	omnivorous	vegetarian	optometrist	head
Female spirit whose wail portends death	banshee	ghoul	background	cow
Referring to a phrase or way of expression natural or peculiar to a particular language	idiomatic	euphemism	imagination	seven
To listen [someone who listens] to other people's conversations	eavesdrop[per]	sleuth	eager	note
Term describing an issue that is open to discussion or modification	negotiable	viable	nobility	pass
Saying little, reserved, uncommunicative	taciturn	withdrawn	tolerance	baby
Very small in size, as in "Gulliver's Travels"	lilliputian	petite	legislation	comedy

Target Definition	Target	Semantic Clue	Phonological Clue	Unrelated Cue
The complete remains of a dead animal, especially at a butcher's	carcass	bones	conjure	wine
Like a beautiful, angelic child	cherubic	saintly	chihuahua	law
State of mind in which one takes or treats things lightly or with a lack of respect	flippant[cy]	sarcastic	flicker	house
Stoppered glass vessel in which spirits are brought to the table	decanter	carafe	depression	smile
Common term for injury to the neck caused by a sudden jerk of the head, for example in a vehicle collision	whiplash	noose	whisper	democracy
Adherent of the view that whether God exists in unknown	agnostic	heathen	adhesive	year
A young goose	gosling	pelican	goblet	beard
A person unnecessarily anxious about their health	hypochondriac	invalid	hemispherical	flute
Medieval forerunner of chemistry	alchemy	transformation	accurate	dish
To [stupefy], bewilder, or confuse	bemuse	stupefy	bouquet	glass
The escape of blood from vessels, including internal as well as external bleeding	hemorrhage	anemia	hurricane	window
Game like hockey but with ball caught by, carried in, and thrown from a net on a stick	lacrosse	badminton	latrine	wife
Lover of books	bibliophile	reader	barbarism	heart
Descriptive term for the type of plant of which the stem is not woody or persistent, and whose leaves and roots are often used for food, medicine, or scent	herbaceous	floral	harmonic	ill
To steam food, particularly meat, slowly in a closed container	braise	sauté	bridge	financial
House of rest for travelers or for the terminally ill, often kept by a religious order	hospice	ambulance	honey	tray
A musical term for a passage performed with gradual decrease in loudness	diminuendo	waning	dissociation	place
Obvious or trite	banal	commonplace	bizarre	judge

APPENDIX H:

TEXT LEVEL POSTTEST QUESTIONS

1. How are the new chapters on the colors and thicknesses of bubbles different from the original book?
    a. they are shorter
    b. they are more difficult
    c. they are not in English
    d. they are the author's favorite

2. The author does not think that there is any loss of precision or accuracy as a result of what?
    a. technical terms of physical science
    b. cumbersome arithmetical treatment of the subject
    c. the avoidance of trigonometry and algebra formulas
    d. difficult chapters added to the book

3. Why has the author seen fit to rearrange, alter, and enlarge his book?
    a. a favorable reception and many translations
    b. a favorable reception and the relevance of bubbles
    c. the favorable opinions of the French, German, and Polish
    d. because there are two tons of bubbles in the world

4. Why does the author use British units?
    a. the unfamiliar metric system would distract the general reader
    b. only highly educated scientists and mathematicians understand these units
    c. Because he thinks British units are archaic
    d. he does not use British units

5. What did the director of the Louvre have to say about the Etruscan vase?
    a. the vase had once been in the collection
    b. no such vase was there
    c. the vase was completely fictitious
    d. the vase was broken at the Louvre long ago

6. Where does Plateau say the only two references to bubbles of any kind are?
    a. **in Ovid and Martial**
    b. on the vase at the Louvre
    c. in the books of different classical authors
    d. Plateau does not provide references to any such amusement

7. The type of student who might find this book useful for preparing for an exam is described as putting more faith in learning formula than what alternative?
    a. reading the chapter on the colors of soap-bubbles
    b. meeting the requirements of any syllabus
    c. learning British methods of measurement
    d. **understanding how they may be recovered when forgotten**

8. In addition to the hints at the end, the author gives:
    a. **guidance for repetition of experiments**
    b. instructions on how to blow marvelous and brilliant bubbles
    c. hints within the text
    d. conveniently located hints in the additional chapters

9. What activity mentioned in the text would demonstrate the forces that this text is about?
    a. Boiling water
    b. Painting
    c. **Pouring tea**
    d. No such activities were mentioned

10. The author assumes that repeating the experiments that he has prepared may:
    a. depend on listening and watching
    b. be dreadfully dull
    c. **should be instructive and interesting**
    d. be informed by somone who knows what they are doing

11. According to the author, what aspects of a bubble might cause someone to wonder how it was produced?
    a. **Its form and colors**
    b. Its motion and luminosity
    c. Its delicate nature and perfection of color
    d. Its size and fragility

12. The author suggests that an experiment is?
    a. not just something to make you wonder

b. not a conjuring trick
c. not shown because of its beauty
d. all of the above

13. According to the author, why do the hairs of the brush cling together when they are underwater?
    a. for a reason we are still unaware of
    b. only because they are wet
    c. they look like they cling together only because the hairs are not separately visible
    d. the hairs do not cling together while submerged under water

14. What does the author claim about vision when you dive in water with your eyes closed?
    a. your vision will be blurrier than if you had dove in with your eyes open
    b. your vision will be better than if you had dove in with your eyes open
    c. you can see the same when you open your eyes as you would if you dove in with them open
    d. your eyelashes prevent you from seeing anything

15. When the author holds the hair brush in the lantern:
    a. the brush catches fire
    b. you can see that the hairs are separate
    c. the heat causes the brush to cling together
    d. the heat dries the wet brush

16. What does the author state is the common belief about why you cannot see properly when you dive into water with your eyes closed?
    a. your vision does not work properly when your eyes are wet
    b. the water gums the eyelashes down over the eyes
    c. you cannot open your eyes while you are under water
    d. your eyelashes do not cling together

17. What does the author claim about water drops with regard to not being mere chance?
    a. the random size
    b. the definite size
    c. the random shape
    d. the uniqueness of each water drop

18. How does the author describe the great drop of water hanging?
    a. similar in size to a water drop about to rupture

b. able to withstand a great pull
c. similar in shape to a water drop about to rupture
d. a heavy liquid

19. What happens when the author allows water to pour in from the pipe?
    a. the bag stretches from the increasing weight of the water
    b. the india-rubber slowly stretches under the increasing weight
    c. the bag reaches a state which indicates that it cannot last much longer
    d. a water-drop falls away

20. According to the author, what did the experiment with the brush first described in his book show?
    a. that hairs do not cling together simply because they are wet
    b. that hairs cling together underwater
    c. water drops have something holding them together
    d. that the surface of water is like an elastic skin

21. In the experiment to demonstrate the ways water behaves as if it had an elastic skin, what evidence illustrates the strength of the skin of the water?
    a. a weight of a quarter ounce placed on the frame is only just sufficient to make the whole thing sink
    b. a weight of four ounces placed on the frame is only just sufficient to make the whole thing sink
    c. a weight of a pound placed on the frame is only just sufficient to make the whole thing sink
    d. none of the above

22. How did the author demonstrate in a more striking way that there is an elastic layer/skin on the surface of water?
    a. by placing paraffin wax in the water
    b. by using a small sieve made out of wire gauze
    c. by explaining how water strider bugs sit on the surface of water
    d. none of the above

23. When the wire frame is placed in the water, why is it able to float in the water?
    a. It does not float
    b. It only floats because of the paper flag
    c. the elastic skin causes it to sink below the surface
    d. the glass globe is large enough

24. How are clean wire and paraffin wax compared in the book?
    a. whether they should be dipped in water

b. whether they affect the skin of water
c. whether they allow a common pin to be put through
d. whether they are wetted by water

25. What is the purpose of the paper in the bottom of the sieve used in the experiment to see ways in which water behaves like it has an elastic skin?
    a. To keep the water in the sieve
    b. To give the wax something to stick to
    c. To make an interesting sound as the water lands in the sieve
    d. To break the fall of the water

26. Does the sieve the author used in an experiment float on water?
    a. yes, the weight of the sieve is not sufficient to stretch the skin of the water through the holes
    b. no, the water passes through the holes
    c. n,o the sieve is too heavy to float
    d. none of the above

27. According to the author, if water has an elastic skin, what should you expect to happen when the wax-coated wire gauze apparatus is dipped in water?
    a. the water will flow easily through the holes
    b. the water should not very readily run through the holes
    c. no water will go through the holes, even when forced
    d. the water will stick to the apparatus

28. Why did the author shake the gauze coated in melted paraffin wax while hot?
    a. to knock the paraffin out of the holes
    b. to spread the paraffin so it can cover all the holes
    c. the author did not shake the gauze while hot
    d. because he burned his hand

29. What is an example the author gives of the power of the elastic skin of water?
    a. spilling water out of a glass
    b. pouring water from a bottle into a tumbler
    c. water leaking from a pipe
    d. pouring water into a narrow-necked bottle

30. According to the author, what do people have to do in the country that is relevant to his discussion of pouring water?
    a. keep water from escaping their water bottles
    b. build iron pipes
    c. carry water from the gutters

d. save money

31. How would it be possible to go to sea in a sieve?
    a. if the sieve is large enough
    b. if the water is not too rough
    c. if the above lines are now realized in every particular
    d. all of the above

32. Why does the water poured into a bottle follow a glass rod held against the edge?
    a. The rod is hollow like a straw and the water flows through it
    b. Because it is coated with wax
    c. The skin of the water along the rod pulls the poured water towards the rod
    d. Static electricity along the rod attracts the water

33. What is the difference between the elastic skin of water and India-rubber mentioned in the book?
    a. elastic skin is perfectly smooth while india-rubber is not
    b. elastic skin is absolutely elastic while india-rubber is not
    c. india rubber is thinner than the elastic skin of water
    d. the elastic skin of water is wet and india-rubber is not

34. What reason does the author give for a phenomenon of water called a capillarity?
    a. the phenomenon occurs when a flow of water is the size of a hair
    b. the phenomenon is demonstrated to be so marked in large tubes
    c. the phenomenon occurs in a tube the size of a hair
    d. the phenomonon only occurs in the capillary vein

35. With a larger tube, how high is the water raised compared to the narrow tube?
    a. double the height
    b. half the height
    c. the same height
    d. four times the height

36. The author states that you can see the water rise in a row of tubes of growing sizes until what?
    a. there is no more water
    b. you run out of tubes
    c. you come to a very large tube
    d. the water will rise indefinitely

37. Why does the author say it will not do to put a narrow tube into quicksilver to show that the level is lower in the tube than in the surrounding vessel?
    - **a. It is not possible to see through quicksilver**
    - b. It is not possible to fit quicksilver into a narrow tube
    - c. the quicksilver would melt the narrow tube
    - d. none of the above

38. In the author's experiment with the level of quicksilver in a tube, what does he say would be the case with water?
    - a. the same
    - b. the level would be lower in the narrow tube
    - **c. the reverse**
    - d. the strength of the elastic skin would be very small

39. What reason does the author give for using quicksilver in his experiment?
    - **a. It does not wet or touch clean glass**
    - b. it is not possible to see through quicksilver
    - c. it can be used with paraffin
    - d. it is thicker than water

40. According to the author, why are the people who live in the country fortunate?
    - a. because they have open space to go out and play
    - b. because they have a better quality of life
    - **c. because they can go to the side of the brook and see the water spiders running on the water without sinking in**
    - d. none of the above

41. What does the author demonstrate by the lines about the centipede?
    - **a. the problem which was solved with the spider**
    - b. water spiders spin nets under water
    - c. the centipede's doubts
    - d. the question asked by the toad

42. According to the author, what does the time it takes for gnat eggs to hatch correspond to?
    - a. silkworm
    - b. caterpillar of a moth
    - c. caterpillar of a butterfly
    - **d. all of the above**

43. How does the author compare the weight of the water spiders and the weight of water?
    a. they are exactly equal
    b. water spiders weigh more than water
    c. water spiders weigh less than water
    d. water spiders weigh much less than water

44. According to the author, how could Mr. H. H. Dixon tell the pressure exerted by the spider on any foot on his scale?
    a. by comparing the bugs' weights
    b. by examining the spiders' feet on the white porcelain dish
    c. by comparing the size of the shadow of the dimples
    d. none of the above

45. How does the author describe the creatures that can be found in water?
    a. small, dark creatures which swim in a curious jerky way
    b. large, dark creatures which swim near the bottom of the water
    c. small, light creatures which swim near the surface
    d. large, light creatures which walk on the ocean floor

46. According to the author, the larva of the gnat:
    a. eat with their breathing tubes
    b. hang from the surface of the water
    c. heavier than the rotten leaves
    d. equal in weight to the downward pull of the water

47. How does the downward pull of larva in water compare to the weight of water needed to fill the small dimple where they hang?
    a. The weight of water is greater than the downward pull
    b. The downward pull is greater than the weight of water
    c. The author would need a better measure to determine how they compare
    d. They are equal

48. The author distinguishes what three parts of a leaf of duckweed when he describes it?
    a. a heel, toe, and finger
    b. a central ridge, heel, and toe
    c. a central ridge, heel, and face
    d. a central ridge, bud, and end

49. In the experiment described in the book, how does the skin of alcohol compare to the skin of water?
    a. it is stronger
    b. they have the same strength
    c. it is weaker
    d. alcohol does not have a skin like water does

50. What does the author compare the game of tug of war to?
    a. the test of weights between alcohol and water
    b. the test of size of drops between alcohol and water
    c. the struggle for area in the white dish between alcohol and water
    d. the test of the strength of skins of alcohol and water

51. How big are the drops of alcohol used in the experiment demonstrating the force of elastic skin of liquids described in the book in comparison to the drops of water?
    a. smaller
    b. bigger
    c. the same size
    d. double the size

52. In the demonstration about the forces of elastic skin of alcohol and water, what happens that causes the author to claim that water is "victorious"?
    a. the water rushes away in all directions
    b. the alcohol sits on the surface of the water
    c. the alcohol penetrates the water
    d. the two form distinct layers

53. Where does the author observe drops of wine?
    a. at the bottom of a wine-glass
    b. spilling out of a wine-glass
    c. on the sides of a wine-glass
    d. mixed with water in a wine-glass

54. What phenomenon does the author use a quote from Proverbs as a reference to?
    a. the evaporation of drops in a glass
    b. the skin of drops of wine
    c. the comparison of the forces of skin of wine drops and water
    d. the movement of drops in a wine-glass

55. How does Dr. Thomson explain the phenomenon the author observes in drops of wine in a wine-glass?
    a. a strong skin
    b. a weak skin
    c. adding water
    d. evaporation

56. What does the author claim about the comparison of skins between ether and water?
    a. the skin of water is weaker than the skin of ether
    b. the skin of ether is weaker than the skin of water
    c. the skins have the same strength
    d. their skins are both weaker than that of wine drops

57. When the author applies clean benzene to a grease-spot, what does he say the greasy benzene will do immediately afterwards?
    a. escape into a fresh rag
    b. remove all the grease
    c. escape into a number of dusters on the ground
    d. run away from the pure benzene ring and heap itself together into the middle

58. Where does the author say you can see a difference between hot and cold grease?
    a. in an open flame
    b. in a burning candle
    c. in lit benzene
    d. in a high-water mark

59. In the author's testing of using benzene on a dress, where did he put the dress?
    a. on the ground outside
    b. in a ring of benzene
    c. over a fresh rag
    d. over dusters outside

60. What happened when the author lit a ring of about a quart of benzene?
    a. It made old dusters look new
    b. it evaporated completely and left everything else untouched
    c. it burned the author's home down
    d. it removed a stain

APPENDIX I:

INFERENCE LEVEL POSTTEST QUESTIONS

1. What do you think the author would say about the use of complicated formulae to teach astrophysics?
    a. **it may have distracted the general audience**
    b. it is helpful to demonstrate the complex relationships
    c. it is not needed for more accuracy and precision
    d. it may give the reader a better understanding of the subject

2. The author assumes which of the following to be true:
    a. readers do not know much about soap bubbles
    b. readers are eager to learn about soap bubbles
    c. **readers are very familiar with soap bubbles**
    d. readers will dislike soap bubbles after reading so much about them

3. A student wants to know what happens when a can is put into a freezer overnight. Which of the following best describes why the author would recommend she "ask Nature a proper question"?
    a. the student will remember it better doing it herself
    b. because other people may have not done the experiment properly
    c. the student will need to replicate your own results to truly understand
    d. **so the student can see the answer for herself**

4. Which of the following is also disproved by the hair brush experiment?
    a. Olympic swimmers do not need to shave their bodies
    b. the sun lightens your hair by separating the hairs as it dries
    c. brushing your hair when it is wet will make it dry
    d. **you should wet a boar's hair brush before brushing wet hair**

5. Based on the India-rubber experiment, what happens to a water droplet after it ruptures?
    a. it falls away because the weight became too heavy
    b. it becomes about half its size smaller because part of its weight was released

   c. **it maintains a similar size and shape even after it ruptures**
   d. any of the above can happen

6. Which of the following best demonstrates the fact that water has an elastic skin based on the globe experiment in the book?
   a. **insects walking on water**
   b. the way anchors sink when dropped at sea
   c. water droplets that are displaced over time in a water bottle
   d. diving into water and making very little splash

7. You pour a cup of water into a bucket with hundreds of 1mm holes at the bottom. Which of the following best describes what would happen?
   a. water cannot seep through when it is poured directly into the basket because the holes are too small
   b. **if water is poured onto a coffee filter, water will seep through if it is displaced somehow**
   c. water will not seep through regardless. One cup is not enough water to force it through the holes
   d. none of the above are accurate statements

8. Which of the following is not necessary in order to successfully got to sea in a sieve?
   a. smooth water conditions
   b. shallow waters
   c. large-sized sieve
   d. **all of the above are necessary**
9. What is the relative strength of the "water skin" in the different sized tubes?
   a. smaller circumference
   b. **it is strongest in the widest tubes because it has the most water to pull up and the largest circumference**
   c. it is strongest in the tallest tube because it has the most height to pull the water
   d. the strength is the same in all tubes

10. Which of the following would be true?
   a. liquid rises higher in wider areas when the liquid is heavier than water
   b. **liquid rises higher in wider areas when the liquid does not actually wet the material**
   c. liquid rises higher in more narrow areas when the liquid is heavier than water
   d. liquid rises higher in more narrow areas when the liquid does not actually wet the material

11. What is an example of an object that is not wetted by water like the spider's foot?
    a. pool rings sinking in a pool
    b. our skin during a bath
    c. the buoys on the lane ropes in pools
    d. none of these objects are wetted by water

12. Differences in atmospheric pressure in water makes which of the following possible?
    a. duckweed connecting end to end
    b. duckweed connecting at the central ridges
    c. duckweed budding underwater
    d. none of the above

13. Which of the following does not accurately describe differences in the weight of two liquids (detergent and alcohol)?
    a. the heavier liquid has smaller drops
    b. the heavier liquid would win a game of "tug of war"
    c. the lighter liquid would have a different water droplet shape
    d. none of the above are true

14. Which of the following is true about the strength of the skin of water?
    a. it can be difficult to alter using vapor
    b. it can be altered easily using vapor
    c. it behaves the same way as wine in a wine glass
    d. it cannot be changed because it will always win the "tug of war"

15. Would going to sea in a sieve be better during the summer or winter months?
    a. summer because the water would have a stronger elastic skin when warmer
    b. winter because the water would have a stronger elastic skin when cold
    c. neither one, spring or fall would be the best in terms of the sea's elastic skin
    d. it would not make a difference

16. Which of the following is most similar to the the intention of this book?
    a. a cookbook that explains the chemistry of baking using formulae
    b. a history book written for people with graduate level historians
    c. an autobiography of a famous mathematician and his experiments
    d. a picture-oriented book about that describes how diesel engines work

17. The historical documentation of soap bubbles could be best described as:
    a. prevalent in works of literature
    b. only prevalent in paintings
    c. commonly documented across many platforms
    d. very rare

18. Which of the following best describes why the author describe experiments in the text?
    a. to demonstrate why well-designed experiments are important
    b. to show that Nature will give you answers if you ask a proper question
    c. to make the book less monotonous
    d. to demonstrate how faulty experiments produce similar results

19. The author describes an experiment where he dips a brush into a cup of water. What conclusion is NOT drawn?
    a. the hairs of a brush will not cling together under water
    b. when the hairs of a brush are dry, they do not cling together
    c. the hairs of a brush cling together after being in water because they are wet
    d. this experiment relates to diving under water with our eyes open

20. Which of the following scenarios is similar to the formation of a water droplet?
    a. a garbage sack that has ripped because of something sharp inside
    b. a garbage sack that has reached its maximum weight
    c. a glass of water that slowly evaporates
    d. a glass of water that sits still

21. Which of the following would be the best way to demonstrate water's "elastic skin"?
    a. pressing a glass globe underwater where bobbles back to the top and is not held down
    b. a heavy glass globe sinking to the bottom of water and remaining there
    c. pressing a glass bowl underwater where it remains until disturbed
    d. a glass bowl slowly sinking to the bottom as you fill it with water slowly

22. Which of the following is the most similar to how water behaves if you poured it in a wax-coated thimble covered with holes?
    a. like a colander for draining
    b. carrying a pail of water with a leak
    c. like a regular cup of water
    d. none of the above

23. Why does water seem to shoot from your fingertips in the shower when you hold your arm at a 45-degree angle away from your side?
    a. it is very lightweight
    b. it is forming a tube around your arm so it will not fall
    c. it gets stuck to the hair on your arm
    d. none of these are correct

24. If you put 10 glass tubes that varied in size in water (1mm to 10mm inches thick), what would be the most likely results?
    a. the water would rise the same height inside and outside all tubes
    b. the water inside the 1-5mm tubes only would be higher than outside the tube
    c. the water would be twice as high in the 4 mm tube compared to the 8mm tube
    d. the water would fill to the same level as outside the tubes

25. Which of the following is required for liquid to be dragged away from narrow spaces?
    a. material not be wetted by the liquid
    b. the liquid be at least as heavy as water
    c. material be a smooth, flat surface
    d. all of the above

26. Which of the following is most similar to the water spider's net that is not wetted?
    a. an air tight tunnel
    b. a fish net under water
    c. a fishing line
    d. all of the above

27. Which of the following would behave most like one item that is wetted and one item that is not wetted?
    a. opposite ends of a magnet, which attract
    b. like-sided ends of a magnet, which repel
    c. hairs of a camel brush
    d. they will behave the same way regardless of being wet

28. How does one liquid win in a game of "tug of war"?
    a. it carries another liquid with it when it moves
    b. it mixes with
    c. another liquid and retains its color
    d. it repels away from another liquid without mixing at all

   e. it does something different than these answer options

29. What result comes from wine on the side of a wine glass?
    a. it remains on the side of the glass longer alcohol evaporates more slowly on the side producing a stronger skin like water
    b. it remains on the side of the glass longer because alcohol evaporates more quickly on the side
    c. it falls down the side of the glass more quickly because alcohol evaporates more slowly on the side producing a weaker skin
    d. it falls down the side of the glass more quickly because alcohol evaporates more quickly on the side producing a weaker skin

30. When you drop some grease on your clean garage floor, what is the first step you should take to remove it?
    a. blot it dry with a rag of greasy benzene
    b. apply pure benzene directly to the spot
    c. apply pure benzene in a circle around the spot
    d. apply a ring of greasy benzene around the spot

BIBLIOGRAPHY

1. Abel, M., & Bäuml, K.-H. T. (2012). Retrieval-induced forgetting, delay, and sleep. *Memory*, *20*(5), 420–428.

2. Anderson, Michael C., and James H. Neely. "Interference and inhibition in memory retrieval." Memory 22 (1996): 586.

3. Ariga, A., & Lleras, A. (2011). Brief and rare mental “breaks” keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition*, *118*(3), 439–443.

4. Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction mind wandering facilitates creative incubation. *Psychological Science*, *23*(10), 1117–1122.

5. Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition*, *20*(4), 1604–1611.

6. Bixler, R., Blanchard, N., Garrison, L., & D’Mello, S. (2015). Automatic Detection of Mind Wandering During Reading Using Gaze and Physiology. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 299–306). ACM.

7. Bixler, R., & D’Mello, S. (2014). Toward Fully Automated Person-Independent Detection of Mind Wandering. In *User Modeling, Adaptation, and Personalization* (pp. 37–48). Switzerland: Springer International Publishing.

8. Bixler, R., & D’Mello, S. (2015). Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Modeling and User-Adapted Interaction*, (pp. 1–36). Springer-Verlag: Berlin Heidelberg.

9. Bixler, R., Kopp, K., & D’Mello, S. (2014). Evaluation of a Personalized Method for Proactive Mind Wandering Reduction. *Proceedings of the 4th International Workshop on Personalization Approaches in Learning Environments (PALE), held*

in conjunction with the 22nd International Conference on User Modeling, Adaptation, and Personalization (UMAP 2014), vol. 1181, pages 33-41.

10. Bjork, R. A., & Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding? *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 567–572.

11. Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444.

12. Blanchard, N., Bixler, R., Joyce, T., & D'Mello, S. (2014). Automated Physiological-Based Detection of Mind Wandering during Learning. In *Intelligent Tutoring Systems* (pp. 55–60). Switzerland: Springer International Publishing.

13. Braboszcz, C., & Delorme, A. (2011). Lost in thoughts: Neural markers of low alertness during mind wandering. *NeuroImage*, *54*(4), 3040–3047.

14. Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *84*(4), 822.

15. Cheyne, J. A., Carriere, J. S., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and Cognition*, *15*(3), 578–592.

16. Cheyne, J. A., Solman, G. J., Carriere, J. S., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, *111*(1), 98–113.

17. Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182.

18. 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.

19. Cirilo, R. K., & Foss, D. J. (1980). Text structure and reading time for sentences. *Journal of Verbal Learning and Verbal Behavior*, *19*(1), 96–109.

20. Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

21. Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

22. Davidson, R. J., Kabat-Zinn, J., Schumacher, J., Rosenkranz, M., Muller, D., Santorelli, S. F., … Sheridan, J. F. (2003). Alterations in brain and immune function produced by mindfulness meditation. *Psychosomatic Medicine*, *65*(4), 564–570.

23. D'Mello, S., Kopp, K., Bixler, R. E., & Bosch, N. (2016). Attending to Attention: Detecting and Combating Mind Wandering during Computerized Reading (pp. 1661–1669). ACM Press.

24. D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*, *70*(5), 377–398.

25. Drummond, J., & Litman, D. (2010). In the zone: Towards detecting student zoning out using supervised machine learning. In Aleven, V., Kay, J., Mostow, J. (eds.) Proceedings of the International Conference on Intelligent Tutoring Systems, (pp. 306–308). Springer, Heidelberg.

26. Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review*, 20(1), 586-592.

27. 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.

28. Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109.

29. Graesser, A. C., Hoffman, N. L., & Clark, L. F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior*, *19*(2), 135–151.

30. Grossman, P., Niemann, L., Schmidt, S., & Walach, H. (2004). Mindfulness-based stress reduction and health benefits: A meta-analysis. *Journal of Psychosomatic Research*, *57*(1), 35–43.

31. Helton, W. S., & Russell, P. N. (2015). Rest is best: the role of rest and task interruptions on vigilance. *Cognition*, *134*, 165–173.

32. Hidi, S. E. (1995). A reexamination of the role of attention in learning from text. *Educational Psychology Review*, *7*(4), 323–350.

33. Jha, A. P., Krompinger, J., & Baime, M. J. (2007). Mindfulness training modifies subsystems of attention. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(2), 109–119.

34. Ju, Y.-J., & Lien, Y.-W. (2016). Better control with less effort: The advantage of using focused-breathing strategy over focused-distraction strategy on thought suppression. *Consciousness and Cognition*, *40*, 9–16.

35. 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.

36. Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47.

37. Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*(6006), 932–932.

38. Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, *29*(2), 133–159.

39. Klinger, E. (1987). Current concerns and disengagement from incentives. In F. Halish and J. Kuhl (Eds.)*, Motivation, intention, and volition,* New York: Springer, 337 – 347 (pp. 337–347).

40. Klinger, E., Barta, S. G., & Maxeiner, M. E. (1980). Motivational correlates of thought content frequency and commitment. *Journal of Personality and Social Psychology*, *39*(6), 1222.

41. Klinger, E., & Cox, W. M. (2004). Motivation and the theory of current concerns. *Handbook of Motivational Counseling: Concepts, Approaches, and Assessment*, 3–27.

42. Kolić-Vrhovec, S., Bajšanski, I., & Rončević Zubković, B. (2011). The role of reading strategies in scientific text comprehension and academic achievement of university students. *Review of Psychology*, *18*(2), 81–90.

43. Kopp, K., Bixler, R., & D’Mello, S. (2014). Identifying Learning Conditions that Minimize Mind Wandering by Modeling Individual Attributes. In *Intelligent Tutoring Systems* (pp. 94–103). Springer International Publishing.

44. Kopp, K., D’Mello, S., & Mills, C. (2015). Influencing the occurrence of mind wandering while reading. *Consciousness and Cognition*, *34*, 52–62.

45. Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. In *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies* (pp. 3–24). Amsterdam, The Netherlands, The Netherlands: IOS Press.

46. Larsen, D. P., Butler, A. C., & Roediger III, H. L. (2009). Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Medical Education*, *43*(12), 1174–1181.

47. MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., … others. (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychological Science*, *21*(6), 829–839.

48. McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*(1), 1–30.

49. 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.

50. McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology. General*, *141*(2), 302.

51. Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, *20*(6), 715–726.

52. Mills, C., & D'Mello, S. (2015). Toward a real-time (Day) dreamcatcher: detecting mind wandering episodes during online reading. In *Proceedings of the 8th International Conference on Educational Data Mining*. Educational Data Mining Society.

53. Mills, C., D'Mello, S. K., & Kopp, K. (2015). The influence of consequence value and text difficulty on affect, attention, and learning while reading instructional texts. *Learning and Instruction*, *40*, 9–20.

54. Mills, C., D'Mello, S., Lehman, B., Bosch, N., Strain, A., & Graesser, A. C. (2013). What Makes Learning Fun? Exploring the Influence of Choice and Difficulty on Mind Wandering and Engagement during Learning. In *Artificial Intelligence in Education* (pp. 71–80).

55. 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.

56. Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science*, 24, 776–781.

57. Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2012). Mindfulness and mind-wandering: finding convergence through opposing constructs. *Emotion*, *12*(3), 442.

58. Murayama, K., Miyatsu, T., Buchli, D., & Storm, B. C. (2014). Forgetting as a consequence of retrieval: a meta-analytic review of retrieval-induced forgetting. *Psychological Bulletin*, *140*(5), 1383.

59. Pavlik, P. I., & Anderson, J. R. (2005). Practice and Forgetting Effects on Vocabulary Memory: An Activation-Based Model of the Spacing Effect. *Cognitive Science*, *29*(4), 559–586.

60. Pettijohn, K. A., Thompson, A. N., Tamplin, A. K., Krawietz, S. A., & Radvansky, G. A. (2016). Event boundaries and memory improvement. *Cognition*, *148*, 136–144.

61. Posner, M. I., & Rothbart, M. K. (2005). Influencing brain networks: implications for education. *Trends in Cognitive Sciences*, *9*(3), 99–103.

62. Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science*, *21*(4), 269–272.

63. Ralph, B. C., Onderwater, K., Thomson, D. R., & Smilek, D. (2016). Disrupting monotony while increasing demand: benefits of rest and intervening tasks on vigilance. Psychological Research, 1-13.

64. Randall, J. G., Oswald, F. L., & Beier, M. E. (2014). Mind-wandering, cognition, and performance: A theory-driven meta-analysis of attention regulation. *Psychological Bulletin*, *140*(6), 1411–1431.

65. Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, *28*(6), 1004–1010.

66. Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye Movements During Mindless Reading. *Psychological Science*, *21*(9), 1300–1310.

67. 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.

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

69. Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255.

70. Roy, M., & Chi, M. T. (2005). The self-explanation principle in multimedia learning. *The Cambridge Handbook of Multimedia Learning*, 271–286.

71. 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.

72. 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.

73. 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.

74. Sibert, J. L., Gokturk, M., & Lavine, R. A. (2000). The reading assistant: eye gaze triggered auditory prompting for reading remediation. In *Proceedings of the 13th annual ACM symposium on User interface software and technology* (pp. 101–107). ACM.

75. Smallwood, J. (2011). Mind-wandering while reading: Attentional decoupling, mindless reading and the cascade model of inattention. *Language and Linguistics Compass*, *5*(2), 63–77.

76. 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.

77. 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.

78. Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: negative moods lead the mind to wander. *Emotion*, *9*(2), 271.

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

80. Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition*, *36*(6), 1144–1150.

81. Smallwood, J., O'Connor, R. C., Sudbery, M. V., & Obonsawin, M. (2007). Mind-wandering and dysphoria. *Cognition and Emotion*, *21*(4), 816–842.

82. Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946.

83. Smith, J. D., & Graham, T. C. (2006). Use of eye movements for video game control. In *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology* (p. 20). ACM.

84. Speca, M., Carlson, L. E., Goodey, E., & Angen, M. (2000). A randomized, wait-list controlled clinical trial: the effect of a mindfulness meditation-based stress reduction program on mood and symptoms of stress in cancer outpatients. *Psychosomatic Medicine*, *62*(5), 613–622.

85. Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236–257.

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

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

88. Tan, T., Zou, H., Chen, C., & Luo, J. (2015). Mind Wandering and the Incubation Effect in Insight Problem Solving. *Creativity Research Journal*, *27*(4), 375–382.

89. VanLehn, K., Jones, R. M., & Chi, M. T. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, *2*(1), 1–59.

90. Vertegaal, R., Slagter, R., Van der Veer, G., & Nijholt, A. (2001). Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 301–308). ACM.

91. Wang, Q., Boccanfuso, L., DiNicola, L., Shic, F., Celebi, F. M., Flink, L., … Kim, E. S. (2015). Interactive eye tracking for gaze strategy modification (pp. 247–250). ACM Press.

92. Whitten, W. B., & Bjork, R. A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior*, *16*(4), 465–478.

93. Wilson, K., & Korn, J. H. (2007). Attention During Lectures: Beyond Ten Minutes. *Teaching of Psychology*, *34*(2), 85–89.

94. Zeidan, F., Johnson, S. K., Diamond, B. J., David, Z., & Goolkasian, P. (2010). Mindfulness meditation improves cognition: evidence of brief mental training. *Consciousness and Cognition*, *19*(2), 597–605.

95. Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 386.

96. Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162.