Do Associations Between Mind Wandering and Learning from Complex Texts Vary by Assessment Depth and Time?

Megan Caruso

University of Colorado Boulder, megan.caruso@colorado.edu

Sidney D'Mello

University of Colorado Boulder, sidney.dmello@colorado.edu

We examined associations between mind wandering – where attention shifts from the task at hand to task-unrelated thoughts – and learning outcomes. Our data consisted of 177 students who self-reported mind wandering while reading five long, connected texts on scientific research methods and completed learning assessments targeting multiple depths of processing (rote, inference, integration) at different timescales (during and after reading each text, after reading all texts, and after a week-long delay). We found that mind wandering negatively predicted measures of factual, text-based (explicit) information and global integration of information across multiple parts of the text, but not measures requiring a local inference on a single sentence. Further, mind wandering only predicted comprehension measures assessed during the reading session and not after a week-long delay. Our findings provide important nuances to the established negative link between mind wandering and learning outcomes, which has predominantly focused on rote comprehension assessed during the learning session itself. Implications for interventions to address mind wandering during learning are discussed.

CCS CONCEPTS **Human-centered computing** → Human computer interaction (HCI); User studies • **Applied Computing** → Law, social and behavioral sciences; Psychology

Additional Keywords and Phrases: Education, mind wandering, user studies, cognitive science

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 INTRODUCTION

Learning from complex texts remains a commonplace learning activity, complementing modern educational technologies such as intelligent tutoring systems, educational games, videos, simulations, and so on. Reading comprehension – particularly reading for understanding and learning – is also a crucially important skill for our society at large; critical in an age of increased skepticism, mistrust, and misinformation [1, 4]. Just as a student must grapple with multiple perspectives in making an informed decision about causes of climate change in a science class [3, 50], a person deciding to get a cancer screening has to confront multiple sources with contradictory recommendations [14, 16, 33]. Thus, in an increasingly information-driven world and workforce, the ability to learn efficiently and effectively from complex texts is critical for success.

Reading outcomes are determined by numerous factors including interactions among the content (the text), the context (the task), and the individual (the reader) [49]. More specifically, outcomes are determined by what the text provides the reader, what the reader brings to the text, and on the affordances of the reading context. Here, we focus on one necessary (but not sufficient) component of successful reading: the ability to sustain attention during reading [8]. Specifically, lapses in attention called "mind wandering" or "zoning out" routinely occur during everyday activities with meta-analytic estimates of approximately 30% during learning from text and video lectures [52]. Mind wandering is also negatively correlated with learning outcomes [52] suggesting that there may be benefits to developing learning interventions that proactively reduce its occurrence [22–24] or that automatically detect and dynamically intervene to reengage attention and mitigate any learning deficiencies [7, 15, 32].

In addition to these technological advances, there is also a need for more basic research on the relationship between mind wandering and learning outcomes. Specifically, as reviewed in Section 1.1, a vast majority (89%) of the studies investigating the relationship between mind wandering and reading comprehension outcomes focus on shallow or rote learning of factual information from the text with learning assessments occurring during reading or immediately after. But reading is a complex task that unfolds across multiple levels: identifying characters, words, and parsing sentences, extracting word-level meaning, representing and connecting textual elements, and incorporating prior knowledge into a constructed mental model of the text [20, 25, 29]. Reading also extends across time. In educational contexts, the goal of reading extends beyond comprehending the text itself, and to retaining the content of what was read over time and transferring that knowledge to new contexts. These processes of comprehension, retention, and transfer depend on the strength and detail of the textual representation while learning from text (henceforth reading).

Accordingly, we examine relationships between mind wandering across levels of comprehension and across different time frames using expository (informational) texts similar to what students encounter for learning. We assess comprehension at multiple depths including knowledge of information explicitly stated in the text (factual), comprehension that connects ideas not implicitly stated in the text (inferencing), and comprehension that involves integration of information across multiple pages and concepts within a text (integration). We also examine comprehension at multiple timescales – during reading a single text, after reading a single text, after reading multiple related texts, and after a week-long delay (retention). To our knowledge, this is the first study to examine relationships between mind wandering and learning and retention from text across different levels of comprehension and time scales.

1.1 Background and Related Work

1.1.1 Reading Comprehension

Reading comprehension is a complex task involving multiple levels of processing which unfold across different depths and across time (Figure 1). According to the prominent construction-integration model, the key processes include: perceptual intake and encoding of the words (surface code), building a literal semantic representation (textbase), and constructing a global situation (mental) model [19, 20]. The first step, the perceptual intake and encoding, involves extracting the words and syntactic structure from the text to form a surface code representation (Figure 1: a, b). The first level of meaning is created by interlinking the representations to form the textbase (construction stage), which also leads to activation of related information in long-term memory (Figure 1: c). Activation spreads until it settles (integration stage), leading to emphasis of concepts that are linked to other concepts in the text or to prior knowledge, and de-emphasis of irrelevant or inconsistent concepts. It is the inferences that are generated during this integration stage that lead to deeper meaning of the text (i.e., the mental model or situation model – Figure 1: d-f)) [12, 13, 19, 28].

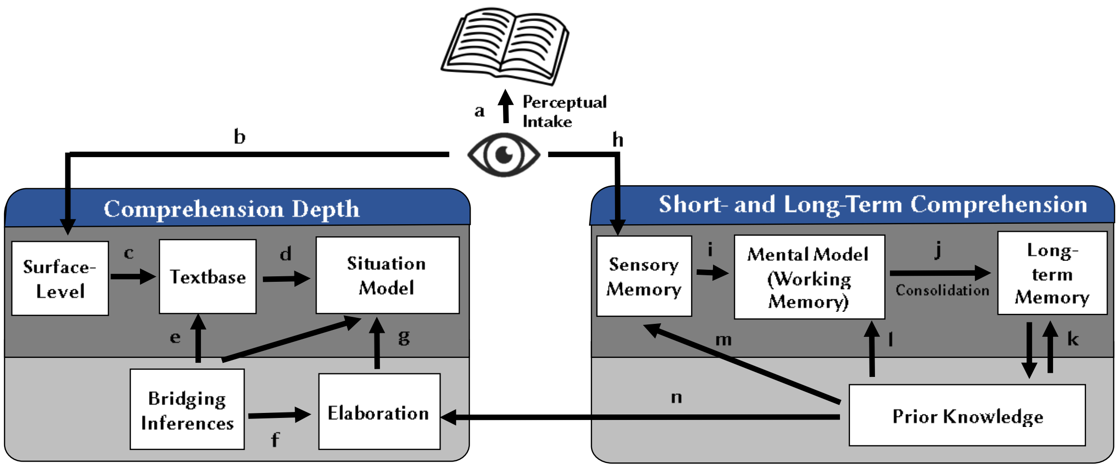


Figure 1. Overview of the process of reading comprehension adapted from Caruso 2022 [5]. Dark grey represents bottom-up processes, and light grey represents top-down processes.

Thus the mental model is affected by both top-down (knowledge-driven) and bottom-up (text-driven) processes [53]; both the text-based information and information from prior knowledge affect the mental model (Figure 1: light grey). An important aspect of comprehension is to go beyond what is explicitly stated in the text to an inter-connected and elaborated representation through inferencing and integration [13, 25, 26]. These inferences can be localized and stem from implicit information extracted from a single sentence [36], or readers can build global inferences across a whole page or multiple pages of text [11, 47], which we refer to as integration. Although there is some debate as which processes are considered inferencing vs. integration, it is generally acknowledged that both are critical to comprehension [27].

Comprehension also unfolds over different time scales (Figure 1: h-n) in addition to different depths. In hundreds of milliseconds a word can be encoded in sensory memory and stored in working memory (Figure 1: h-i), where representations can be maintained for several seconds, prior to being consolidated into long-term memory. Prior knowledge from long-term-memory in turn affects how information from a text is represented, processed, and remembered (Figure 1: k-n) via inferencing and integration [27]. Additionally, subsequent reading can interfere with memory traces of what is currently being read before it is consolidated into long term memory [37, 54], thereby affecting the information recalled in the following days, weeks, and beyond. Memory retrieval is itself not a passive process. Retrieving memories for a text strengthens encoding, creates new retrieval processes, and alters existing memories [17, 40]. Indeed, reading for understanding is not a simple passive process, but a complex active event which unfolds across different time scales.

1.1.2 *Mind Wandering During Reading Comprehension*

Mind wandering, also called task-unrelated thought, or more commonly "zoning out" and "daydreaming", is an ubiquitous phenomenon [48]. It can be defined as *involuntary* shifts in attention from the task at hand to task-unrelated thoughts [8], but also see [35, 46] for debates on mind wandering with intentionality. Estimates of how often mind wandering occurs range from 30% to 50% [18, 21, 52] depending on the task, context, and individual. Although there are some potential benefits of mind wandering on activities that require prospection and divergent thinking [34], it is consistently negatively associated with performance for tasks that require sustained attention, such as learning. Indeed, Wong et. al (2022)'s recent meta-analysis of 71 studies from 5890 participants on mind wandering during educational activities reported that mind wandering occurred on average 30% of the time irrespective of activity types (video learning, reading, in-person learning) [52]. Further, mind wandering had a negative correlation of r-0.27 on assessments of learning, which replicates results of prior meta-analyses on mind wandering and performance across a range of learning and cognitive tasks [8, 39].

It is prudent to consider why mind wandering is negatively associated with learning outcomes. The perceptual decoupling hypothesis [43] suggests that during mind wandering attention is decoupled from the external stimulus (e.g., the text) and instead is directed inward (e.g., what's for dinner) resulting in encoding failures [43, 45] and a weakened mental model, which in turn leads to lower comprehension outcomes. Beyond perceptual decoupling [43], which in the case of reading affects the lower-order processes such as the surface and textbase representations, higher-order reading processes (integration and inferencing) can also be decoupled during mind wandering. Specifically, Mills et al introduced the idea of cognitive decoupling where readers who mind wandered failed to adjust cognitive resources with respect to text difficulty [31]. Perceptual and cognitive decoupling have negative downstream effects: according to the cascade model, deficits in the early processes at lower levels of comprehension *cascade* to deeper levels, resulting in an impaired mental model.

1.2 Contribution, Novelty, and Hypotheses

Research indicates that mind wandering negatively correlates with reading comprehension, but most studies (63 out of 71 samples in Wong's (2022) meta-analysis [52]) assessed factual memory of the text (rote comprehension) rather than deeper levels of comprehension involving inferencing and integration. Of the studies that focused on inference-level comprehension, correlations with mind wandering ranged from -.02 to -.40, a very wide range. Although Wong et al. do not provide a clear definition of what they consider inference-level assessments, examination of the primary studies reveals that these assessments mainly test for integration across multiple sentences [2, 30, 38, 42]. Overall, the small number of studies focusing on deeper levels of comprehension is surprising considering an important goal of learning is to construct meaningful representations connecting different textual elements and going beyond what is stated in the text by elaborating based on prior knowledge. Indeed, Wong et al. end with a call for action regarding the importance of providing clearer definitions of assessment type and more studies examining inferencing and integration with text comprehension and mind wandering [52]. Additionally, most studies assess comprehension during reading or immediately after, but never at a longer delay, which is surprising as reading is primarily for the purpose of retention and transfer. The strength of the relationship between mind wandering and reading comprehension at different depths and across time is currently unexamined.

We address these critical gaps by investigating the relationship between mind wandering and learning with assessments targeting (1) different depths of comprehension and (2) different time onsets between the initial encoding and subsequent learning assessments. We utilized a large dataset of 177 participants who read five long, connected expository texts while also indicating the occurrence of mind wandering during reading. We assessed rote (factual), inference (local integration), and deep (global integration) comprehension at four time points: during reading, after reading each text, after reading all texts, and at least seven days after reading.

We hypothesize (Hypothesis 1) that mind wandering will be negatively associated with comprehension based on the principles of perceptual and cognitive decoupling discussed earlier and consistent with numerous prior works. However, our critical hypothesis (Hypothesis 2) is that the strength of this relationship will be moderated by the depth and time of the comprehension assessment. Specifically, because comprehension where deficits at lower levels (textbase) cascade and affect subsequent deeper-level processing (inference and integration), as proposed by the cascade model of inattention, mind wandering will be more negatively associated with deeper assessments (inference/integration compared to shallower levels of comprehension (rote assessments – Hypothesis 2a). A competing hypothesis, however, is that because a literal encoding of the text is needed to answer rote assessments, any nonencoded information due to mind wandering cannot be subsequently retrieved from memory. Conversely, prior knowledge and heuristic processing can still aid the readers to fill in gaps for assessments targeting deeper levels of comprehension. Thus, the correlation between mind wandering and comprehension will be weaker for deeper comprehension compared to rote assessments (competing Hypothesis 2b). Lastly, we do not provide a hypothesis with respect to associations between mind wandering and comprehension assessed at different time scales, instead consider this to be an exploratory research question.

2 METHODS

2.1 Data Source

2.1.1 Participants

Data was collected as part of a larger study investigating neurophysiology during reading comprehension. Only aspects applicable to current study are discussed here. Participants (N= 177, age 23 ± 7 years, 68.3% female, 76% white, 10% Asian, 7 % Hispanic, 4% other, <1% Black or African American, <1% Native Hawaiian or Pacific Islander) were students from a large public University in the Western US. Each participant was compensated $20/hour for the duration of the study (143 ± 26 minutes), and an additional $10 for completing the follow-up survey emailed a week after their participation day. Compensation was via Amazon gift cards. All procedures were approved by the Institutional Review Board (IRB) and all participants provided written informed consent.

2.1.2 Materials & Procedure

The main study occurred in a research lab with eye tracking and neurophysiological sensing not discussed further here.

Prior Knowledge. Before beginning the reading task, participants completed a prior knowledge assessment consisting of six concepts and their definitions, all of which relate to scientific methods subsequently covered in the reading. Participants had to match the six randomly ordered definitions with the six randomly ordered concept names (Placebo, Dependent Variable, Confound, Control Group, External Validity, and A priori). Correct answers were summed, with the highest score of six (all terms matched correctly) and the lowest score of zero (no items matched correctly).

Texts. The reading material consisted of five expository texts on the topic of behavioral research methods: Bias, Hypotheses, Causal Claims, Validity, and Variables. Each text was around 1000 words split into 10 pages (screens of text). The five texts were presented in a random order to each participant (while the 10 pages within a given text were always presented in the same order for that text). The texts were specifically designed to be suitable yet slightly advanced for college students with a mean Flesch-Kincaid grade level of 13.2 [10]. Participants then read through the texts one page at a time, using a key to advance to the next page, but were unable to move back to already-read pages. The average time spent reading a text was 5.5 minutes (1.8 SD), with a total average reading time of 27 minutes (8.1 SD).

Mind Wandering. We used the probe-caught method to track mind wandering; this method has been extensively validated in the literature [44, 51]. Specifically, before beginning the reading task, participants were given the definition of mind wandering and informed they would be probed to self-report whether they were mind wandering or not while reading. Specifically, we used the following instructions.

"At some points during reading, you may realize that you were not thinking about what you are actually reading, but were thinking about something else altogether. This is called "zoning out". We're interested in such zone-outs and you will sometimes be asked to report if you have zoned out while reading. Please be as honest as possible about reporting zoning out. It is perfectly natural to zone out while reading. Responding that you were zoning out will in no way affect your progress in this study, so please be completely honest with your reports".

After reading the instructions, participants were given the following multiple-choice question to verify they understood mind wandering and how to identify it. If an incorrect response was chosen, then the participant was informed they answered incorrectly, were once again shown the definition of zoning out, and were given another try until they answered correctly.

Which below is an example of zoning out?

A) When you accidentally choose the wrong answer in a text
B) When you fall asleep while reading a boring book
C) When you find yourself thinking about something other than the task at hand *(correct)*
D) When you struggle to understand the text

Probes occurred twice per text on pseudorandom pages, when participants pressed the key to advance from page P to page P + 1. Specifically, participants were prompted with a question asking if they were mind wandering on the previous page (P) and were given "yes" or "no" answers to choose from. Each participant received ten mind wandering probes over the course of reading, two per text. The average MW rate across participants in the study was 47.0% (24.8%), which was higher than the 30% estimate from the Wong meta-analysis [52], but likely due to the length of the reading session.

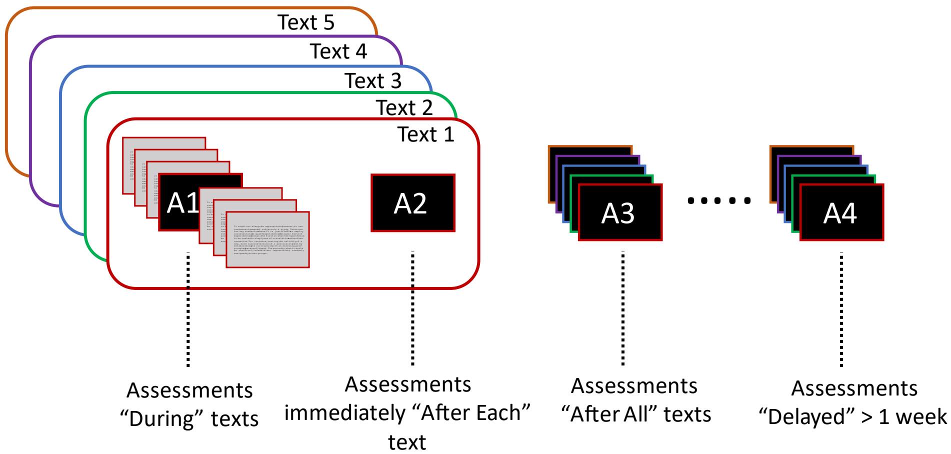
Comprehension/Learning Assessments. Three types of assessments were designed to assess different levels of comprehension. They consisted of: 1) rote questions that consisted of a four-item multiple choice targeting factual knowledge explicitly presented in the text; 2) inference questions consisting of a statement reflecting either a true or false inference that can be made by extracting implicit information from a specific single sentence on the page; and 3) integration questions consisting of a four-item multiple choice question targeting deep, integrative knowledge across multiple pages. See Table 1 for examples of each question type. Comprehension questions were developed and refined by researchers with the guidance of an instructor of a behavioral research methods undergraduate course. We used crowdsourcing via Mechanical Turk to iteratively modify and calibrate the questions to the desired difficulty: items with greater than 20% accuracy or less than 80% accuracy (reflecting potential floor and ceiling effects, respectfully) were either discarded or adapted until mean accuracies for all comprehension questions were within this range.

Table 1. Examples of the three different comprehension questions. Text is taken directly from the reading were possible and summarized and in cases where it cannot fit into the table.

Question Type	Text From the Reading	Answers
Rote	"When the *person asking the question was behind a wall* so that Hans could not see these small gestures, the horse could not pick up on what the correct answer might be and therefore could not answer questions correctly"	Please select the best answer to the question: How did the experimenters prevent the horse from seeing the person asking questions? *A) By placing the horse on the other side of a wall (correct)* B) By fitting blinkers to the horse C) By putting a veil over the questioner's face D) By making the horse and questioner face in opposite directions *The text states that there was a wall between the horse and the person, it does not mention any of the other options.*
Inference	"The second [scenario where it is justifiable to employ correlational design] is when it would be unethical, infeasible or impossible to randomly assign subjects to groups"	Is the statement true or false? "Random assignment would not be possible for a study with gender as the independent variable." *A) True (correct)* *B)* False *Gender, by nature, is not something that you can assign to a person and therefore random assignment of gender in an experiment is impossible.* Note that this information is not contained directly in the text but must be inferred.
Integration	Page 2: Developing a new measure that might correlate highly with intelligence, how this relates to convergent and divergent validity. … Page 6-7: External validity and how most research subjects are college students … Page 8: Defines content validity and gives an example.	Please selected the best answer to the question: Which of the following could be a reason why a measure of intelligence lacked content validity? *A) All of the questions relate to mathematical and verbal skills, but none of the questions relate to spatial skills (correct)* B) It lacks face validity C) It correlates strongly with an existing measure of working memory D) Most of the data was collected using white, affluent college students between the ages of 18 and 24 *This question requires the reader to apply the definition of content validity from the text to a new example.*

Comprehension assessments were administered at four different time points, but not all question types were administered at each time point due to limits on the total item pool (See Figure 2): (A1) "During" text occurring immediately after reading the corresponding page, all questions types were administered; (A2) "After Each" occurring after an individual text was read, only Rote and Inference questions were administered; (A3) "After All" occurring once all five texts were completed, all question types were administered; (A4) "Delay" occurring a minimum of seven days after the reading session, only Rote and Inference were administered. Questions appeared in the order: Rote, Inference, Integration to reflect the increasing levels of processing. The mind wandering probes always occurred immediately prior to the "During" assessments to allow participants to be as accurate as possible in responding to the mind wandering probes. Each comprehension assessment type was administered twice per text at the given time point.

The selection of pages to trigger comprehension questions was pseudorandom with some basic constraints (i.e., no questions on page 1 or on adjacent pages) but the specific questions administered to a participant were pseudorandomly selected from a main question pool without repetition for the "During", "After Each", and "After All" time points. Thus, each participant received a different set of questions at different pages for these three time points. However, all participants received the same delayed assessment with some overlap in pages assessed at earlier time points (while the questions were randomized between participants for the first three time points, Delay questions were not since there was a single survey sent out to all participants a week later).



	A1	A2	A3	A4
Rote	X	X	X	X
Inference	X	X	X	X
Integration	X		X	

Figure 2. Four different time points and the types of comprehension questions asked at each. Black boxes indicate assessments within the pages of a text (grey boxes), and colors indicate the five individual texts (and subsequently the corresponding text of the assessment questions).

2.2 Analytical Approach

We used multilevel ordinal regression at the text-level using the *clmm* function from the R package *ordinal* [6], with the number of correct answers per text as the outcome variable (answered no questions correctly: 0, answered one out of the two questions correctly: 1, answered both questions correctly: 2). Participant and the Text (categorical variable indicating one of the five texts) were included as random intercepts (more complex random effect structures resulted in convergence errors). Text-level mind wandering rate (ordinal, answered no to both mind wandering probes while reading the text: 0, answered yes to one mind wandering probes: 1, answered yes to both mind wandering probes: 2) was included in a three-way interaction (which also includes all two-way interactions) with comprehension type (categorical for rote, inference, and integration), and time (a factor indicating "During", "After Each", "After All", or "Delay", with "During" as the reference variable). We also controlled for two variables: the order each text was presented to the participant (fatigue effects) and pre-test scores (to account for prior knowledge). Overall, the equation takes the form:

$$\text{Correct} \sim \text{Mind Wandering} * \text{Question Type} * \text{Time} + \text{Text Order} + \text{Prior Knowledge} + (\text{MW}|\text{Participant}) + (1|\text{Text})$$

Results for the ordinal mixed effects models are given in odds ratios, where a significant odds ratio above 1 indicates a positive predictor, and less than 1 indicates a negative predictor. Odds ratios are an effect size metric, so a value of 1.22 means that one unit increase in that predictor would make it 1.22 times more likely to result in an increase in comprehension scores. We probed significant trends with the *emtrends* function from the *emmeans* package in R [41]. A false discovery rate (FDR) was used to adjust for multiple testing for each set of trends. Additionally, a two-tailed $p < 0.05$ criterion was used.

3 RESULTS

3.1 Descriptive Statistics and Correlations

We first examined subject-level correlations (Table 2) between the comprehension assessments, mind wandering, and prior knowledge. We calculated the proportion correct for each comprehension assessment for each participant, as well as the average mind wandering rate for each participant. Comprehension scores ranged from 1.13 (out of 2) to 1.51 reflecting accuracies of 57% to 7% with a mean of 69%. On average, integration scores (1.14 or 57%) were lower than rote and inference scores, which were on par (1.41 or 70%). Overall, however, there do not appear to be floor- or ceiling-effects.

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Mind Wandering (0-2)	0.94	0.50											
2. Prior Knowledge (0-6)	4.16	1.15	-0.21**										
Rote (0-2)													
3. "During"	1.51	0.36	-0.26**	0.40**									
4. "After Each"	1.42	0.34	-0.21**	0.28**	0.45**								
5. "After All"	1.35	0.42	-0.18*	0.42**	0.49**	0.50**							
6. "Delay"	1.34	0.41	-0.22**	0.41**	0.44**	0.35**	0.46**						
Inference (0-2)													
7. "During"	1.50	0.29	-0.21**	0.29**	0.43**	0.35**	0.45**	0.28**					
8. "After Each"	1.44	0.27	-0.24**	0.19**	0.17*	0.19*	0.23**	0.24**	0.24**				
9. "After All"	1.46	0.28	-0.18*	0.20*	0.16*	0.20**	0.27**	0.32**	0.10	0.14			
10. "Delay"	1.26	0.30	0.00	0.15	0.08	-0.02	0.13	0.07	0.08	0.01	-0.01		
Integration (0-2)													
11. "During"	1.13	0.37	-0.22**	0.36**	0.50**	0.43**	0.50**	0.35**	0.39**	0.21**	0.23**	0.09	
12. "After Each"	1.15	0.37	-0.31**	0.29**	0.39**	0.22**	0.37**	0.37**	0.23**	0.16*	0.15*	0.14	0.21**

Table 2. Means (M), Standard Deviations (SD) and Pearson Correlations between Mind Wandering, Prior Knowledge, and Comprehension assessments, * indicates $p < 0.05$. ** indicates $p < 0.01$. Ranges are shown next to each variable heading. Numbers are given to each label on the right and correspond to the numbers on the top of the table.

Correlations between comprehension assessments were a mean of 0.22, suggesting the different assessments captured related but unique variance. As expected, mind wandering was negatively correlated with all comprehension assessments (mean $r = -.20$) except for inference items assessed at Delay, which incidentally was also weakly correlated with the other comprehension measures. Prior knowledge was negatively correlated with mind wandering ($r = -0.21$) and all comprehension measures (mean $r = 0.29$), indicating the importance of controlling for it.

3.2 Model 1: Rote, Inference, and Integration Comprehension During and After Reading

To compare the effects of mind wandering on the three levels of comprehension depth we examined the two time points at which all three assessments were administered: "During" and "After All" (n = 5,211 cases). Of particular interest is the three-way mind wandering × comprehension depth × assessment time interaction along with the two-way mind wandering × comprehension depth and mind wandering × assessment time interactions. Of these, there was only a significant two-way interaction between mind wandering and comprehension depth. Consequently, we examined the relationship between mind wandering (independent variable) and comprehension outcomes (dependent variables) for different comprehension types (moderator). As shown in Figure 3, mind wandering was a significant negative predictor of integration (B = -.16 [-.24, -.08]) and rote (B = -.16 [-.24, -.08] comprehension), but not inference (B = -.04 [-.12, .04]) comprehension. In terms of comparisons of the slopes, both integration and rote comprehension were on par (p = .983) and significantly (p = .034) higher than inference comprehension.

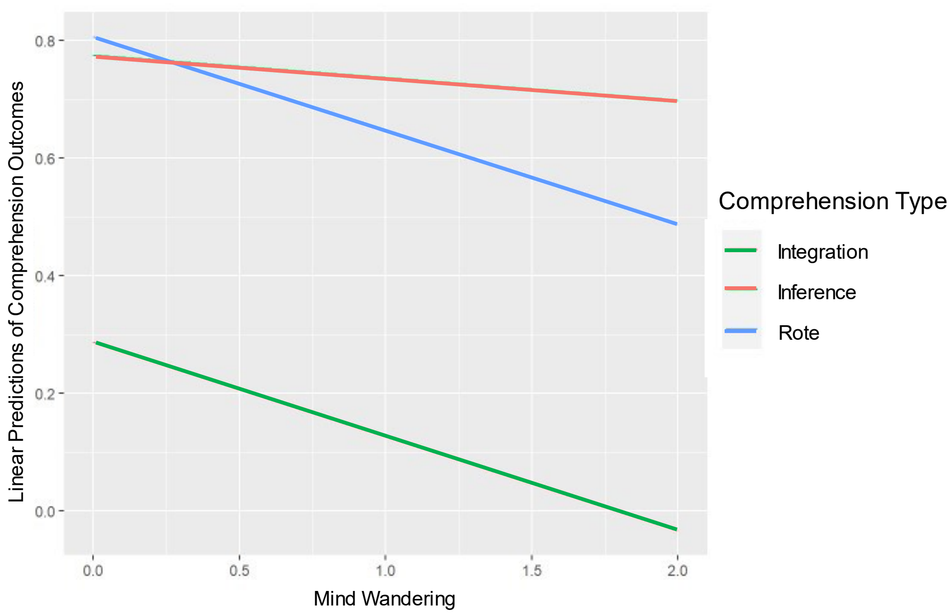


Figure 3. Trends for correct as mind wandering increases for each comprehension type: Rote, Inference, and Integration, over all four assessment timings. Y axis are in log odds.

In terms of the random effects, the between-participant variance ($\tau_{00\text{ParticipantID}}$) was .10 and the between-text variance ($\tau_{00\text{Text}}$) was .08. The intra-class correlation (ICC) measures the proportion of variance explained by the nesting factors, in this case ParticipantID and Text and was .14, suggesting that 14% of the variance in comprehension outcomes was explained by the ParticipantID and Text grouping. The marginal R^2 indicates the proportion of variance explained by the fixed effects (10.6%), and the conditional R^2 is the proportion of the variance explained by the fixed and random effects (23%).

3.3 Model 2: Inference and Rote Comprehension During, After Each, After All, and Delayed

To compare the effects of mind wandering over additional time points, we also fit a model that focused on Rote and Inference comprehension, which were administered at all four time points: "During", "After Each", "After All", and "Delayed" (n = 6765 cases). Once again, there were significant two-way interactions for mind wandering × comprehension, but there was also a significant mind wandering × time interaction, but no significant three-way interaction. As shown in Figure 4, mind wandering was a significant negative predictor of rote (B = -.15 [-.21, -.09]) but not inference (B = -.04 [-.10, .02]) comprehension (rote had a significantly steeper slope than inference, p = 0.002), replicating results from Model 1 at these two additional time points.

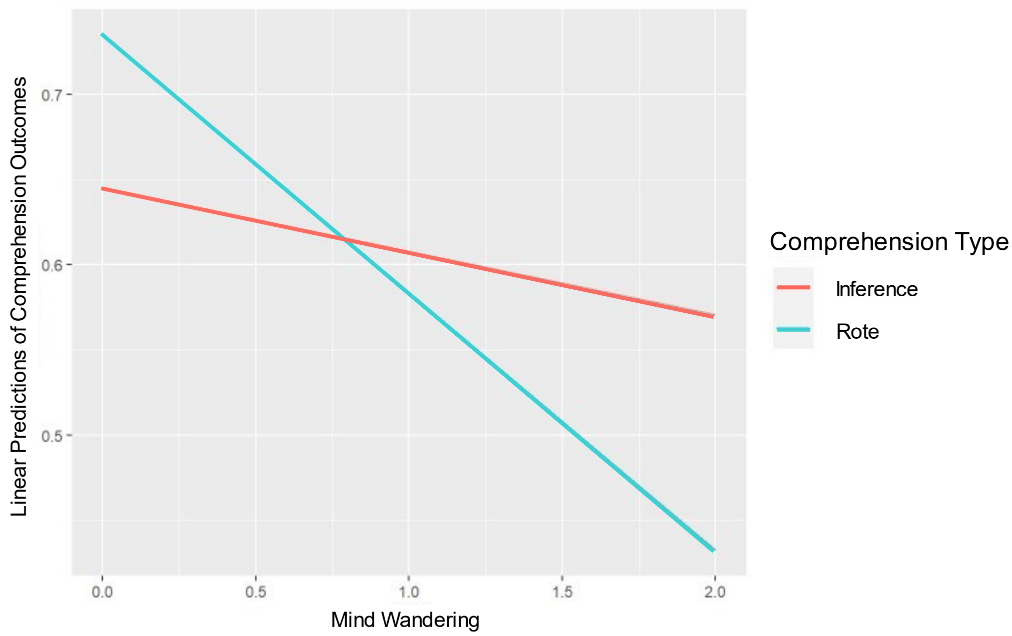


Figure 4. Trends for correct comprehension as mind wandering increases for each comprehension type: Rote, Inference, over all four assessment timings. Y axis are in log odds.

Next, we also examined the relationship between mind wandering (independent variable), and comprehension outcomes (dependent variable) for different times (the moderator). As shown in Figure 5, mind wandering was a significant negative predictor for comprehension assessed during (B = -.13 [-.21, -.05]) reading, after reading each text (B = -.14 [-.22, -.06]), and after reading all texts (B = -.08 [-.17, -.01]), but not after a week-long delay (B = -.02 [-.10, .07]).

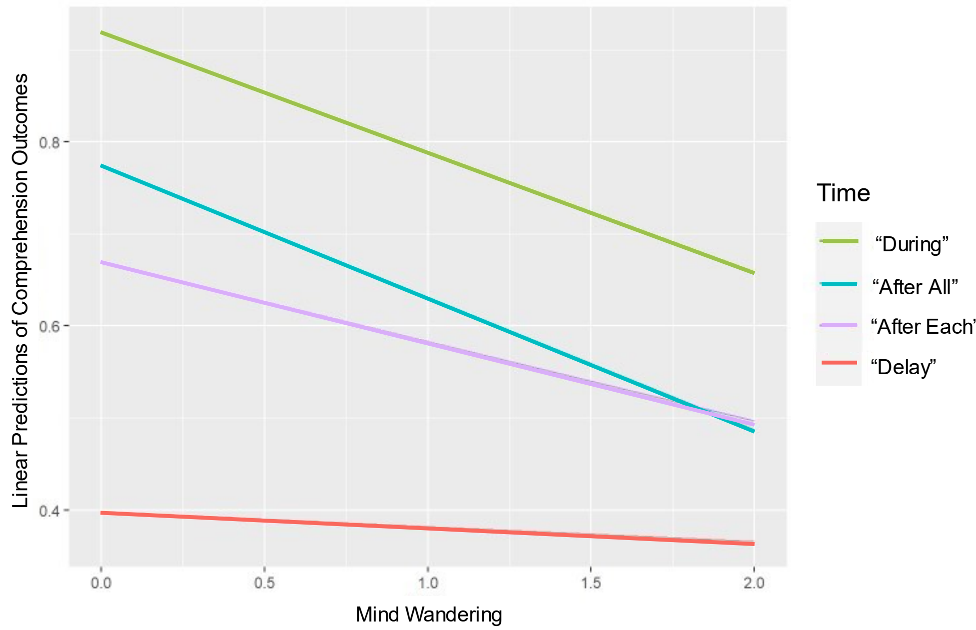


Figure 5. Model 2: Trends for correct comprehension as mind wandering increases for each time point at which comprehension was assessed. Y axis are in log odds.

4 DISCUSSION

Our study examined the effect of mind wandering on comprehension outcomes on a large dataset of students who read five long, connected, informational texts. We tested the readers' ability to effectively extract information from and connect ideas in the text at varying degrees of depth and delay (including up to a week later). In what follows, we summarize our main findings, implications, along with limitations and directions for future work.

4.1 Main Findings

We found that mind wandering occurred on 47% of pages on which it was measured, which is a higher than the average meta-analytic estimate of 30% that Wong et al found when reviewing 71 studies [52]. The somewhat higher rate might be attributed to the fact that the learning task was long (spanning more than an hour) and challenging, but might also be attributable to the lab setting (see below).

Overall, we found that mind wandering negatively correlated with comprehension and learning outcomes, with a mean *r* of -.20, consistent with the Wong et al. (2022) meta-analysis. However, this relationship was moderated by both the type and time of the comprehension assessments. Specifically, mind wandering was a significant negative predictor of rote and integration comprehension with similar effect sizes. This is inconsistent with Hypothesis 2a which would predict that mind wandering would have stronger effects on deeper levels of comprehension. Our competing hypothesis (Hypothesis 2b) predicted that mind wandering might have less of an effect on deeper level comprehension due to readers ability to utilize prior knowledge and heuristic processing, but this was also not supported. Instead, only inference level assessments had differential (i.e., null effects) with mind wandering. Recall that the inference assessments targeted implicit information from a single sentence on a page, instead of across multiple sentences or pages (integration). Our working hypothesis is that learners could have relied on their constructed mental models, prior knowledge, or reasoning to answer these questions despite mind wandering. But future work is needed to further explore this surprising null effect.

Lastly, we found that mind wandering negatively predicted comprehension assessed during reading, after each text, and after reading all texts, but not for comprehension assessed a week later. This suggests that mind wandering is primarily associated with deficiencies with initial encoding and mental model construction, but not with subsequent processes of consolidation into long-term memory. This finding highlights the need for more research which examines comprehension not only at different depths, but also different time scales, which is largely overlooked in the literature [52].

4.2 Implications

This research can improve current efforts to develop computational models of mind wandering and comprehension. Various models have been developed to detect mind wandering using eye gaze while reading [9, 15, 32], and some have even been applied in real-time. Hutt et al, for example, used gaze to detect mind wandering and launch interventions to reengage students when mind wandering was detected [15]. Similarly, Mills et al prompted readers to self-explain missed concepts if mind wandering was detected and re-read if needed [32]. The design of optimal interventions in both these examples can benefit from a deeper understanding of the relationship between mind wandering and different types of comprehension outcomes. For instance, the lack of a relationship between mind wandering and local inferencing compared to rote and integration assessments suggests that interventions might instead be tailored to improve these later outcomes.

Many systems designed to mitigate mind wandering and its outcomes also fail to account for long-term comprehension, even though this is a key goal when such systems are deployed in the classroom. If more is known regarding this relationship, the online prompts to the reader can be optimized to improve retention, not solely comprehension during reading. Indeed, the Mills et al study above found benefits of real-time detection and mitigation of mind wandering on comprehension assessed a week after reading rather than immediately after reading.

4.3 Limitations & Future Work

Our study has some limitations. First, the lab setting of our experiment might have changed participants behavior compared to how they would act in more ecologically valid settings. Specifically, participants were wearing several neural and physiological sensors as well as masks to adhere to COVID safety procedures. This might have resulted in discomfort and unnatural behaviors while reading, which in turn might lead to somewhat higher rates of mind wandering. Second, the study did not administer all question types at all time points due to a limited item pool, which restricted our analysis across time for different depths of comprehension. Specifically, integration questions were only provided during and immediately after reading and not after each text or at the delayed time point. Lastly, the participants only read one set of texts, so it is not known if these results would generalize to new texts or different types of comprehension assessments. Future research is needed to replicate results with different texts and in different contexts and with expanded assessments.

5 CONCLUSION

Learning from texts is a complex cognitive process critical to most educational activities and is a process that unfolds over different depths and different time scales. Attending to the text is key when encoding and integrating information, yet a major gap in the literature is that the vast majority of studies examining the relationship between mind wandering and comprehension outcomes focus on rote (factual) memory of the text assessed during or immediately after reading. Given that deeper levels of processing and retention over time are highly valued learning outcomes, our study was the first to include an expansive set of comprehension assessments at different depths and time. Our findings indicate that the relationship between mind wandering and comprehension was complex in that it was moderated by both the depth of comprehension assessment (no effects for local inferencing but similar effects for rote and global integration assessments) and time course of assessment (no effects for assessments after a week-long delay). These non-intuitive effects have important implications for theories of the role of mind wandering during reading comprehension and for interventions that aim to improve learning from text by detecting and combatting mind wandering as it occurs.

6 ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation (DRL 1920510). The opinions expressed are those of the authors and do not represent views of the funding agency.

References

[1] Alexander, P.A. and The Disciplined Reading and Learning Research Laboratory 2012. Reading Into the Future: Competence for the 21st Century. *Educational Psychologist*. 47, 4 (Oct. 2012), 259–280. DOI:https://doi.org/10.1080/00461520.2012.722511.

[2] Bixler, R. and D'Mello, S. 2016. Automatic gaze-based user-independent detection of mind wandering during computerized reading. *User Modeling and User-Adapted Interaction*. 26, 1 (Mar. 2016), 33–68. DOI:https://doi.org/10.1007/s11257-015-9167-1.

[3] Bråten, I., Strømsø, H.I. and Salmerón, L. 2011. Trust and mistrust when students read multiple information sources about climate change. *Learning and Instruction*. 21, 2 (Apr. 2011), 180–192. DOI:https://doi.org/10.1016/j.learninstruc.2010.02.002.

[4] Britt, M.A., Richter, T. and Rouet, J.-F. 2014. Scientific Literacy: The Role of Goal-Directed Reading and Evaluation in Understanding Scientific Information. *Educational Psychologist*. 49, 2 (Apr. 2014), 104–122. DOI:https://doi.org/10.1080/00461520.2014.916217.

[5] Caruso, M., C, Ace Peacock, Southwell, R., Guojing Zhou and D'Mello, S. 2022. Going Deep and Far: Gaze-based Models Predict Multiple Depths of Comprehension During and One Week Following Reading. (Jul. 2022). DOI:https://doi.org/10.5281/ZENODO.6852998.

[6] Christensen, R. 2019. *ordinal---Regression Models for Ordinal Data*.

[7] D'Mello, S.K., Mills, C., Bixler, R. and Bosch, N. 2017. Zone out no more: Mitigating mind wandering during computerized reading. *Proceedings of the 10th International Conference on Educational Data Mining (EDM 2017).* (2017).

[8] D'Mello, S.K. and Mills, C.S. 2021. Mind wandering during reading: An interdisciplinary and integrative review of psychological, computing, and intervention research and theory. *Language and Linguistics Compass*. 15, 4 (2021), e12412. DOI:https://doi.org/10.1111/lnc3.12412.

[9] Faber, M., Link to external site, this link will open in a new window, Krasich, K., Bixler, R.E., Brockmole, J.R. and D'Mello, S.K. 2020. The eye–mind wandering link: Identifying gaze indices of mind wandering across tasks. *Journal of Experimental Psychology: Human Perception and Performance*. 46, 10 (Oct. 2020), 1201–1221. DOI:http://dx.doi.org/10.1037/xhp0000743.

[10] Flesch, R. 1948. A new readability yardstick. *Journal of Applied Psychology*. 32, 3 (1948), 221–233. DOI:https://doi.org/10.1037/h0057532.

[11] Garrod, S. and Sanford, A. 1977. Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*. 16, 1 (Feb. 1977), 77–90. DOI:https://doi.org/10.1016/S0022-5371(77)80009-1.

[12] Graesser, A.C. and Hemphill, D. 1991. Question answering in the context of scientific mechanisms. *Journal of Memory and Language*. 30, 2 (Apr. 1991), 186–209. DOI:https://doi.org/10.1016/0749-596X(91)90003-3.

[13] Graesser, A.C., Singer, M. and Trabasso, T. 1994. Constructing inferences during narrative text comprehension. *Psychological Review*. 101, 3 (Jul. 1994), 371–395. DOI:https://doi.org/10.1037/0033-295X.101.3.371.

[14] Han, P.K.J., Klein, W.M.P., Lehman, T., Killam, B., Massett, H. and Freedman, A.N. 2011. Communication of Uncertainty Regarding Individualized Cancer Risk Estimates: Effects and Influential Factors. *Medical Decision Making*. 31, 2 (Mar. 2011), 354–366. DOI:https://doi.org/10.1177/0272989X10371830.

[15] Hutt, S., Krasich, K., R. Brockmole, J. and K. D'Mello, S. 2021. Breaking out of the Lab: Mitigating Mind Wandering with Gaze-Based Attention-Aware Technology in Classrooms. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama Japan, May 2021), 1–14.

[16] Jensen, J.D., Krakow, M., John, K.K. and Liu, M. 2013. Against conventional wisdom: when the public, the media, and medical practice collide. *BMC Medical Informatics and Decision Making*. 13, 3 (Dec. 2013), S4. DOI:https://doi.org/10.1186/1472-6947-13-S3-S4.

[17] Karpicke, J.D. and Janell R., B. 2011. Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*. 331, 6018 (2011), 772--775.

[18] Killingsworth, M.A. and Gilbert, D.T. 2010. A Wandering Mind Is an Unhappy Mind. *Science*. 330, 6006 (Nov. 2010), 932–932. DOI:https://doi.org/10.1126/science.1192439.

[19] Kintsch, W. 1998. *Comprehension: a paradigm for cognition*. Cambridge University Press.

[20] Kintsch, W. 1994. Text comprehension, memory, and learning. *American Psychologist*. 49, 4 (Apr. 1994), 294–303. DOI:https://doi.org/10.1037/0003-066X.49.4.294.

[21] Klinger, E. and Cox, W.M. 1987. Dimensions of Thought Flow in Everyday Life. *Imagination, Cognition and Personality*. 7, 2 (Oct. 1987), 105–128. DOI:https://doi.org/10.2190/7K24-G343-MTQW-115V.

[22] Kopp, K., Bixler, R. and D'Mello, S. 2014. Identifying Learning Conditions that Minimize Mind Wandering by Modeling Individual Attributes. *Intelligent Tutoring Systems* (Cham, 2014), 94–103.

[23] Kopp, K., Mills, C. and D'Mello, S. 2016. Mind wandering during film comprehension: The role of prior knowledge and situational interest. *Psychonomic Bulletin & Review*. 23, 3 (Jun. 2016), 842–848. DOI:https://doi.org/10.3758/s13423-015-0936-y.

[24] Krasich, K., Hutt, S., Mills, C., Spann, C.A., Brockmole, J.R. and D'Mello, S.K. 2018. "Mind" TS: Testing a Brief Mindfulness Intervention with an Intelligent Tutoring System. *Artificial Intelligence in Education* (Cham, 2018), 176–181.

[25] McCarthy, K.S. and McNamara, D.S. 2021. The Multidimensional Knowledge in Text Comprehension framework. *Educational Psychologist*. 56, 3 (Jul. 2021), 196–214. DOI:https://doi.org/10.1080/00461520.2021.1872379.

[26] McKoon, G. and Ratcliff, R. 1992. Inference during reading. *Psychological Review*. 99, 3 (1992), 440–466. DOI:https://doi.org/10.1037/0033-295X.99.3.440.

[27] McNamara, D.S. 2021. If Integration Is the Keystone of Comprehension: Inferencing Is the Key. *Discourse Processes*. 58, 1 (Jan. 2021), 86–91. DOI:https://doi.org/10.1080/0163853X.2020.1788323.

[28] McNamara, D.S. and Magliano, J. 2009. Chapter 9 Toward a Comprehensive Model of Comprehension. *Psychology of Learning and Motivation*. Academic Press. 297–384.

[29] Millis, K.K., Long, D.L., Magliano, J. and Wiemer, K. 2018. *Deep comprehension: Multi-disciplinary approaches to understanding, enhancing, and measuring comprehension*. Routledge.

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

[31] Mills, C., Graesser, A., Risko, E.F. and D'Mello, S.K. 2017. Cognitive coupling during reading. *Journal of Experimental Psychology: General*. 146, 6 (2017), 872–883. DOI:https://doi.org/10.1037/xge0000309.

[32] Mills, C., Gregg, J., Bixler, R. and D'Mello, S.K. 2021. Eye-Mind reader: an intelligent reading interface that promotes long-term comprehension by detecting and responding to mind wandering. *Human–Computer Interaction*. 36, 4 (Jul. 2021), 306–332. DOI:https://doi.org/10.1080/07370024.2020.1716762.

[33] Montgomery, M. 2010. Uncertainty During Breast Diagnostic Evaluation: State of the Science. *Oncology Nursing Forum*. 37, 1 (Jan. 2010), 77–83. DOI:https://doi.org/10.1188/10.ONF.77-83.

[34] Mooneyham, B.W. and Schooler, J.W. 2013. The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*. 67, (2013), 11–18. DOI:https://doi.org/10.1037/a0031569.

[35] Murray, S. and Krasich, K. 2022. Can the mind wander intentionally? *Mind & Language*. 37, 3 (2022), 432–443. DOI:https://doi.org/10.1111/mila.12332.

[36] Myers, J.L. and Duffy, S.A. 1990. Causal Inferences and Text Memory. *Psychology of Learning and Motivation*. Elsevier. 159–173.

[37] Nadel, L., Hupbach, A., Gomez, R. and Newman-Smith, K. 2012. Memory formation, consolidation and transformation. *Neuroscience & Biobehavioral Reviews*. 36, 7 (Aug. 2012), 1640–1645. DOI:https://doi.org/10.1016/j.neubiorev.2012.03.001.

[38] Naylor, J.S. and Sanchez, C.A. 2018. Can reading time predict mind wandering in expository text? *Applied Cognitive Psychology*. 32, 2 (2018), 278–284. DOI:https://doi.org/10.1002/acp.3393.

[39] Randall, J.G., Oswald, F.L. and Beier, M.E. 2014. Mind-wandering, cognition, and performance: A theory-driven meta-analysis of attention regulation. *Psychological Bulletin*. 140, (2014), 1411–1431. DOI:https://doi.org/10.1037/a0037428.

[40] Roediger, H.L. and Butler, A.C. 2011. The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*. 15, 1 (Jan. 2011), 20–27. DOI:https://doi.org/10.1016/j.tics.2010.09.003.

[41] Russell, L. 2022. *emmeans: Estimated Marginal Means, aka Least-Squares Means*.

[42] Sanders, J.G., Wang, H.-T., Schooler, J. and Smallwood, J. 2017. Can I Get me out of my Head? Exploring Strategies for Controlling the Self-Referential Aspects of the Mind-Wandering State during Reading. *Quarterly Journal of Experimental Psychology*. 70, 6 (Jun. 2017), 1053–1062. DOI:https://doi.org/10.1080/17470218.2016.1216573.

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

[44] Schubert, A.-L., Frischkorn, G.T. and Rummel, J. 2020. The validity of the online thought-probing procedure of mind wandering is not threatened by variations of probe rate and probe framing. *Psychological Research*. 84, 7 (Oct. 2020), 1846–1856. DOI:https://doi.org/10.1007/s00426-019-01194-2.

[45] Seibert, P.S. and Ellis, H.C. 1991. Irrelevant thoughts, emotional mood states, and cognitive task performance. *Memory & Cognition*. 19, 5 (Sep. 1991), 507–513. DOI:https://doi.org/10.3758/BF03199574.

[46] Seli, P., Risko, E.F. and Smilek, D. 2016. On the Necessity of Distinguishing Between Unintentional and Intentional Mind Wandering. *Psychological Science*. 27, 5 (May 2016), 685–691. DOI:https://doi.org/10.1177/0956797616634068.

[47] Singer, M., Graesser, A.C. and Trabasso, T. 1994. Minimal or Global Inference during Reading. *Journal of Memory and Language*. 33, 4 (Aug. 1994), 421–441. DOI:https://doi.org/10.1006/jmla.1994.1020.

[48] Smallwood, J. and Schooler, J.W. 2015. The Science of Mind Wandering : Empirically Navigating the Stream of Consciousness. *Annual Review of Psychology*. (Jan. 2015), 487–518.

[49] Snow, C. 2002. *Reading for Understanding. Towards an R&D Program in Reading Comprehension*. Office of Educational Research and Improvement.

[50] Strømsø, H.I., Bråten, I. and Britt, M.A. 2010. Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*. 20, 3 (Jun. 2010), 192–204. DOI:https://doi.org/10.1016/j.learninstruc.2009.02.001.

[51] Varao-Sousa, T.L. and Kingstone, A. 2019. Are mind wandering rates an artifact of the probe-caught method? Using self-caught mind wandering in the classroom to test, and reject, this possibility. *Behavior Research Methods*. 51, 1 (Feb. 2019), 235–242. DOI:https://doi.org/10.3758/s13428-018-1073-0.

[52] Wong, A.Y., Smith, S.L., McGrath, C.A., Flynn, L.E. and Mills, C. 2022. Task-unrelated thought during educational activities: A meta-analysis of its occurrence and relationship with learning. *Contemporary Educational Psychology*. 71, (Oct. 2022), 102098. DOI:https://doi.org/10.1016/j.cedpsych.2022.102098.

[53] Woolley, G. 2011. *Reading Comprehension*. Springer Netherlands.

[54] Yeari, M. 2017. The role of working memory in inference generation during reading comprehension_ Retention, (re)activation, or suppression of verbal information? *Learning and Individual Differences*. 56, (2017), 1–12. DOI:https://doi.org/10.1016/j.lindif.2017.04.002.