

What Makes Learning Fun? Exploring the Influence of Choice and Difficulty on Mind Wandering and Engagement during Learning

Caitlin Mills¹, Sidney D'Mello^{1,2}, Blair Lehman³, Nigel Bosch², Amber Strain³, and Art Graesser³

Departments of Psychology¹ and Computer Science², University of Notre Dame
Notre Dame, IN 46556, USA
{cmills4, sdmello, pbosch}@nd.edu

Department of Psychology and Institute for Intelligent Systems³, University of Memphis,
Memphis, TN 38152, USA
{dchuncey, balehman, a-graesser}@memphis.edu

Abstract. Maintaining learner engagement is critical for all types of learning technologies. This study investigated how choice over a learning topic and the difficulty of the materials influenced mind wandering, engagement, and learning during a computerized learning task. 59 participants were randomly assigned to a text difficulty and choice condition (i.e., self-selected or experimenter-selected topic) and measures of mind wandering and engagement were collected during learning. Participants who studied the difficult version of the texts reported significantly higher rates of mind wandering ($d = .41$) and lower arousal both during ($d = .52$) and after the learning session ($d = .48$). Mind wandering and arousal were not affected by choice. However, participants who were assigned to study the topic they selected reported significantly more positive valence during ($d = .57$) but not after learning. These participants also scored substantially higher on a subsequent knowledge test ($d = 1.27$). These results suggest that choice and text difficulty differentially impact mind wandering, engagement, and learning and provide important considerations for the design of ITSs and serious games with a reading component.

Keywords: engagement, mind wandering, reading, serious games, affect

1 Introduction

Keeping learners attentive and engaged has long been an important challenge for computerized learning systems. Although learners might begin a session with some enthusiasm and involvement, engagement wanes as time passes [1–3] and learners start to disengage by zoning out or engaging in unproductive, off-task behaviors [4–6]. These types of behaviors have been linked to negligible learning, lowered interest, and attrition in academic contexts [6–8]. The problem of diminished or outright disengagement during a learning session threatens the effectiveness of educational tech-

nologies because engagement is a necessary (but not sufficient) condition for learning, particularly at deeper levels. Therefore, advances in uncovering and detecting the factors that trigger disengagement are sorely needed.

Engagement is a complex meta-construct with behavioral, affective, and cognitive components that vary both situationally and dispositionally [9]. Effort and task persistence constitute some of the behavioral components of engagement [9], while the affective components include valence, arousal, and discrete emotions like interest and curiosity. The cognitive components of engagement include attention, concentration, and the use of learning strategies. There have been an increasing number of studies that focus on the behavioral and affective components of engagement [10–12], yet very little attention has been given to some of the cognitive components.

One such component is the phenomenon of mind wandering (or zoning out or day-dreaming). Mind wandering is the attentional shift away from processing external, task-related information towards the processing of internal, task-unrelated information [13]. Mind wandering is detrimental to a range of educational activities as reviewed by [14]. This is because active comprehension involves extracting information from the learning environment and aligning this information with existing mental models that are ultimately consolidated into long term memory structures [15–18]. A coupling between external information (task) and internal representations (existing mental model) is essential for meaningful comprehension of the material. Mind wandering signals a breakdown in this coupling process [19–20].

To date, very little research in the AIED and ITS communities have been devoted to the study of mind wandering. One notable exception is a study by [4] that focused on using acoustic-prosodic and lexical features to detect self-reported instances of zoning out during a spoken learning session. Hence, the present paper consists of some basic research to identify the factors that influence engagement and mind wandering during a computerized learning task.

One important factor that might play a role in maintaining engagement during learning sessions is the difficulty of the material. For example, [21] reported that mind wandering was more frequent when participants read difficult texts compared to easy texts and that mind wandering also had a more negative impact on comprehension for the difficult texts. However, this study used narrative texts, so there is the question of whether these findings generalize to learning from academic texts.

Another factor that might impact engagement is the perception of choice over the learning material. The control-value theory of emotion posits that learners' appraisals of subjective control and value about an activity predict the emotions that will arise during a learning session [3, 22]. Engagement is hypothesized to be higher when learners have control and some autonomy over the learning task [23–24]. One pioneering study by [25] provided some evidence to support this claim. They gave learners choices over non-instructional components of a serious game (e.g., character icons and names). Learners who were given choices liked the system better, wanted more time with the system, and performed better on a math test. More recently, [26] found that when children had control over an interactive storybook, they showed more interest and less dramatic declines in attention, compared to when adults were in control. Another study by [27] found that more interest was reported when learners chose the

order in which texts were presented. Interest, in turn, influenced affect, learning, and persistence.

The studies discussed above have focused on the influence of choice and difficulty on promoting engagement. However, these factors have been studied in isolation, so there is the question of whether these factors interact to influence engagement. For example, are difficult topics more engaging when learners perceive a choice over the topics? In line with these questions, the goal of the present research was to investigate how text difficulty and perceived choice affect engagement and learning during a computerized learning task consisting of reading instructional texts. We focused on text reading because students arguably spend more time studying from textbooks than other learning activities and reading is often considered to be non-interactive and boring. Reading is therefore an excellent context to investigate engagement.

The texts used in the present study were modified versions of materials from a serious game called *Operation ARIES!* [28]. *Operation ARIES!* teaches scientific critical thinking through a series of modules, including reading about core concepts from an online textbook and having conversations with animated pedagogical agents. We focused on the reading portion, because it lacks interactivity and it is solely up to the learner to maintain attention during reading in order to learn the material.

The current experiment had a 2×2 (text difficulty \times perceived choice) between subjects design. For the difficulty manipulation, participants received an easy or difficult version of a scientific reasoning text. For the choice manipulation, participants were given a choice of two text titles, and either received the text they selected to read (self-selected) or the text they did not select (experimenter-selected). Engagement was measured in two ways: (1) self-reported levels of valence and arousal (affective component) and (2) mind wandering reports via auditory probes, which is a standard way to track mind wandering [13, 29]. We focus on three research questions: (1) What is the rate of mind wandering during a computerized learning task?, (2) What is the impact of perceived choice and text difficulty on mind wandering, valence, and arousal?, and (3) Do perceived choice and text difficulty affect text comprehension?

2 Method

2.1 Participants and Design

There were 59 participants recruited from Amazon's Mechanical Turk™ (AMT). AMT allows individuals to receive monetary compensation for completing Human Intelligence Tasks online. Participation was limited to native English speakers at least 18 years of age. The mean age was 38.4 years old ($SD = 12.3$). On average, the study lasted 22 minutes and participants were compensated \$1.75. Past research suggests AMT is a reliable and valid source for collecting experimental data [30-31]. There are also some advantages to using AMT with respect to diversity, at least when compared to typical undergraduate samples used in many research studies.

The experiment had a 2×2 between subjects design in which choice (self-selected vs. experimenter-selected) and text difficulty (easy vs. difficult) were randomly assigned. Details on these manipulations are given below.

2.2 Materials

Text Manipulations. The experimental texts were adapted from two texts about research methods used in the serious game, *Operation ARIES!* [28]. Both texts focused on a research methods concept: (1) the dependent variable and (2) making causal claims. Texts began with a case study that demonstrated how the respective concept applies to real world situations and followed with explanations and examples demonstrating uses for the concept.

Easy and difficult versions were created for each text by manipulating the two texts on the following dimensions: narrativity, sentence length, word frequency, syntactic simplicity, and referential cohesion. These were identified by [32] as the textual features that contribute to text difficulty and conceptual clarity. Easy versions were created to be more narrative, with shorter sentences and fewer low frequency words. They were also made more cohesive by replacing ambiguous pronouns with proper nouns. Difficult texts had longer, more complex sentences with more low frequency words. Both versions, however, had the same conceptual content and were approximately 1500 words.

Significant differences in text difficulty were assessed by comparing easy and difficult texts via three measures: (1) Flesch-Kincaid Grade Level (FKGL), (2) Coh-Metrix (a text-analysis software) indices of difficulty [33], and (3) subjective human ratings. First, we ensured that the FKGL were at least two grade levels different. Easy texts were at grade 9 and difficult texts were grade 11. Second, we looked at a more systematic assessment of difficulty based on the Coh-Metrix indices of difficulty (narrativity, referential cohesion, deep cohesion, and syntactic simplicity). Higher values of each index indicate that a text is easier to read. Easy and difficult texts were significantly different based on these four indices in the expected direction (average $p < .05$). Finally, we completed a pilot study to make sure that humans perceived the texts to differ in levels of difficulty. Humans rated the difficult texts to be significantly more difficult after reading ($d = .93$), $p < .05$. There were also no differences between the two texts (e.g., easy dependent variable text compared to easy causal claims text) among these three dimensions.

Learning Measures. Learning was measured through multiple-choice deep reasoning questions (nine questions per text). These questions were developed in adherence to the Graesser-Person question asking taxonomy [34] specifically targeting logical, causal, or goal-oriented reasoning. Each participant received a three-question pretest and a six-question posttest, which corresponded to the specific text they read.

2.3 Procedure

After filling out an electronic consent form, participants completed a pretest that consisted of three deep reasoning questions to assess prior knowledge, followed by instructions for the self-paced learning task. Self-paced reading was adopted for this task to eliminate any pressures from time constraints.

The choice feedback manipulation occurred before participants began reading the text. First, participants were presented with two different headlines (one for each text) and were asked to choose which one they would like to read. The headlines were: (dependent variable) “Are you being controlled by subliminal messages hidden in plain sight?” and (making causal claims) “Wipe that tired expression off your face! This new energy pill is bound to put some pep in your step!”

After selecting a headline, participants were immediately given feedback to indicate whether or not they would be given their selected text to read. Participants were randomly assigned to either receive the text they selected (self-selected) or the text they did not select (experimenter-selected). Participants who received the self-selected text were given the message, “Good news for you! You’ll read the text you wanted to read!” Alternatively, participants who received the experimenter-selected text received the following message: “Unfortunately, you’ll be reading the text you did not choose. Too bad.” This feedback manipulation explicitly informed participants about whether or not their headline selection influenced the text they received.

Prior to engaging in the self-paced reading, participants were informed that an auditory probe (i.e., a beep) would periodically sound during reading. At the time of the probe, they were instructed to indicate whether or not they were currently mind wandering by hitting “Y” (yes) or “N” (no) on the keyboard. The following description of mind wandering, taken from previous studies [13, 21], was provided to the participants to aid in distinguishing mind wandering episodes: “At some point during reading, you may realize you have no idea what you just read. Not only were you not thinking about the text, you were thinking about something else altogether.” A total of ten auditory mind wandering probes were inserted in each text. The probes corresponded to pages that contained content that was relevant to the learning measure. A sentence-by-sentence reading paradigm allowed probes to be located at more precisely controlled content locations across easy and difficult texts.

In addition to the mind wandering probes, participants were asked to report levels of valence and arousal at three separate points: before, during (the middle), and after reading the text. Valence was measured on a 6-point scale from 1 (very negative) to 6 (very positive). Arousal was measured with a similar scale ranging from 1 (very sleepy) to 6 (very active). Finally, a six-item posttest was completed after the learning session.

3 Results and Discussion

3.1 Mind Wandering

There were a total of 590 mind wandering probes across the 59 participants. The distribution of mind wandering proportions was non-normal, so non-parametric statistics were used for significance testing involving this variable. The mean proportion of probes to which participants responded “yes” was .354, indicating that mind wandering occurred approximately one third of the time participants were probed. Indeed, this finding reveals that participants reported mind wandering over 30% of the time

during this computerized learning task, highlighting an important concern for the prevalence of this phenomenon.

There is a question of whether perceived choice and text difficulty influenced levels of mind wandering. A Mann-Whitney U Test revealed that there was significantly more mind wandering in the difficult condition (33.7%) compared to the easy condition (20.3%), $Z = -1.95$, $p = .051$. Perceived choice, however, did not impact rates of mind wandering, $p = .654$ (see Table 1 for descriptive statistics on mind wandering).

3.2 Valence and Arousal

Participants reported their valence and arousal levels at three different points: before, during, and after reading. Delta valence and arousal scores were computed by subtracting before scores from *during* and *after* scores (delta during and after valence and arousal). These two delta measures were used in order to control for participants' baseline valence and arousal levels. Table 1 provides descriptive statistics for the delta valence and arousal measures.

Univariate analyses of variance (ANOVAs) revealed a main effect of perceived choice on delta valence during reading, $F(1, 55) = 4.52$, $p = .038$, partial $\eta^2 = .076$. Participants who read the self-selected text reported negligible changes in valence *during* reading ($M = .029$, $SD = .674$) compared to the participants who read the experimenter-assigned text ($M = -.360$, $SD = .700$). However, there was no perceived choice effect for the change in valence *after* reading, $F(1, 55) = 1.10$, $p = .300$.

Interestingly, the main effect of text difficulty yielded quite different patterns for valence and arousal. Whereas perceived choice influenced valence, text difficulty impacted arousal. There was a marginally significant main effect of text difficulty on delta arousal *during* reading, $F(1, 55) = 3.74$, $p = .058$, partial $\eta^2 = .064$. Participants who read the difficult text ($M = -.233$, $SD = .897$) showed a larger drop in arousal in the middle of the reading compared to the participants who read an easy text; arousal actually increased for those participants who read an easy text ($M = .172$, $SD = .658$). Similarly, there was a marginally significant effect of text difficulty on delta arousal *after* reading, $F(1, 55) = 3.40$, $p = .071$, partial $\eta^2 = .058$. There was a larger drop in arousal for participants who read a difficult text ($M = -.300$, $SD = 1.06$) compared to an easy text ($M = .138$, $SD = .743$) *after* reading. However, text difficulty did not impact valence either *during* or *after* reading.

These findings indicate that perceived choice and text difficulty differentially impacted valence and arousal. Perceived choice increased valence *during* reading ($d = .57$), whereas text difficulty was associated with a decrease in arousal *during* ($d = .52$) and *after* reading ($d = .48$). There were no interactions of perceived choice and text difficulty with respect to valence and arousal.

It is also worth noting that delta valence and arousal *during* and *after* reading were negatively correlated with mind wandering. Non-parametric correlations indicated that mind wandering was negatively correlated with delta arousal *during* ($r_s = -.256$, $p = .050$) and *after* ($r_s = -.329$, $p = .011$) reading. Similarly, delta valence *during* ($r_s = -.100$, $p = .453$) and *after* ($r_s = -.317$, $p = .015$) reading were also negatively correlated with mind wandering.

3.3 Text Comprehension

Participants' performance on the pretest and posttest were computed as the proportion of items answered correctly. In order to control for prior knowledge, corrected learning gains were calculated from these scores as: $(\text{Posttest} - \text{Pretest}) / (1 - \text{Pretest})$. A univariate ANOVA indicated that participants who read the self-selected text ($M = .473$, $SD = .300$) had significantly higher learning gains compared to those who read the experimenter-assigned text ($M = -.153$, $SD = .628$), $F(1, 54) = 24.6$, $p < .001$. Text difficulty did not impact learning gains nor did it interact with perceived choice.

This finding further supports the control-value theory of emotions and previous work on autonomy and choice. Those participants who felt as if they had a choice in the learning material performed significantly better on the comprehension test compared to those who did not perceive a choice ($d = 1.27$). A heightened sense of subjective value might be inherent in the ability to choose learning materials, leading to deeper engagement and learning.

Table 1. Descriptive Statistics (M) for Mind Wandering, Valence, Arousal, and Corrected Learning Gains based on Text Difficulty and Perceived Choice.

	Text Difficulty			Perceived Choice		
	Easy	Diff	d	Self	Exp	d
Mind Wandering (Proportion)	.244	.461	.413	.357	.351	.012
Valence During	-.172	-.100	.102	.029	-.360	.567
Valence After	-.310	-.267	.046	-.177	-.440	.271
Arousal During	.172	-.233	-.515	-.088	.040	-.154
Arousal After	.138	-.300	-.480	-.059	.886	.064
Corrected Learning Gains	.238	.192	-.083	.473	-.153	1.27

4 General Discussion

Sustaining students' engagement over time in any ITS or serious game is still an important concern. This paper provides insight for how two factors, namely text difficulty and perceived choice, impact engagement during a non-interactive reading task. Results suggest giving learners choices about their learning material might be a simple way for systems to advantageously maintain engagement, specifically capitalizing on the control aspect in the control-value theory of emotions [22]. One idea is to focus on the choice of certain materials over others (e.g., choose between these two texts), rather than the choice of the order of materials (e.g., choose the order you will read these texts). Specifically, systems could employ this technique and facilitate engage-

ment by creating the illusion of choice. The selection options can be highly ambiguous (more or less interchangeable), such that the target learning material can be presented regardless of the option that was selected. For example, if the target learning material is a text on the scientific method, two headlines can be presented that both could feasibly align with the text. Regardless of which headline the participant selects, the same text could then be presented, giving the participant a greater sense of control by having made a choice.

The results of the present study also indicated that the difficult texts were associated with lowered engagement levels. Therefore, it is important to design learning materials that will adequately challenge learners, without being so difficult that attention cannot be sustained. Texts that are too difficult might induce lower engagement, as well as increase the risk of attentional lapses from the external environment, which is obviously undesirable for the duration of a learning session. The importance of difficulty of the learning material is not a novel idea [16-17]; however, this study is the first evaluation of how text difficulty and perceived choice affect mind wandering in a computerized learning task with academic texts.

It is important to note the limitations of this study. For example, a longer text would allow us to track how these factors affect engagement over a longer period of time. Another limitation is that we did not measure any individual differences of topic and situational interest, which have been previously related to choice manipulations [27]. Understanding individual differences, such as these, might improve models of engagement by incorporating how learners' traits interact with factors from the learning environment. Also, although previous research found a negative relationship between mind wandering and learning [21], we did not replicate this finding. This warrants further testing with different sets of academic texts over different time domains, as this learning session was relatively short (about 1500 words).

Lastly, since our study was conducted online, we were unable to collect any eye tracking or physiological measurements of engagement. These additional measures could aid in developing a more fine-grained model of mind wandering and engagement. Combining task factors like the ones in this experiment with physiological measures and eye tracking can be an initial step towards predicting when a learner begins to mind wander and/or disengage from a text. Interventions can then be put into place in order to restore attentional focus to the current learning task.

Acknowledgment. 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.

References

1. Csikszentmihalyi, M.: Flow: The psychology of optimal performance. Cambridge University Press, New York (1990).

2. D’Mello, S., Graesser, A.: The half-life of cognitive-affective states during complex learning. *Cognition & Emotion*. 25, 1299–1308 (2011).
3. Pekrun, R., Goetz, T., Daniels, L., Stupnisky, R., Perry, R.: Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*. 102, 531 (2010).
4. Drummond, J., & Litman, D. In the zone: Towards detecting student zoning out using supervised machine learning. In: Alevan, V. Kay, J., Mostow, J., (Eds.) *Intelligent Tutoring Systems*. Part II, vol 6095. Springer, Berlin/Heidelberg, 306–309 (2010).
5. Forbes-Riley, K., & Litman, D. When does disengagement correlate with learning in spoken dialog computer tutoring? In 15th International Conference on *Artificial Intelligence in Education*, 81–89 (2011).
6. Mann, S., Robinson, A.: Boredom in the lecture theatre: An investigation into the contributors, moderators and outcomes of boredom amongst university students. *British Educational Research Journal*. 35, 243–258 (2009).
7. Baker, R., D’Mello, S., Rodrigo, M., Graesser, A.: Better to be frustrated than bored: The incidence, persistence, and impact of learners’ cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*. 68, 223–241 (2010).
8. Larson, R., Richards, M.: Boredom in the middle school years: Blaming schools versus blaming students. *American Journal of Education*. 99, 418–443 (1991).
9. Fredricks, J., Blumenfeld, P., Paris, A.: School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*. 74, 59–109 (2004).
10. Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Meheranian, H., Fisher, D., et al.: Repairing disengagement with non-invasive interventions. *Frontiers in Artificial Intelligence and Applications*. 158, 195 (2007).
11. D’Mello, S., Chipman, P., & Graesser, A.: Posture as a predictor of learner’s affective engagement. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, 905–910 (2007).
12. Joseph, E.: Engagement tracing: Using response times to model student disengagement. *Artificial Intelligence in Education: Supporting Learning Through Intelligent and Socially Informed Technology*. 125, 88 (2005).
13. Smallwood, J., Schooler, J.: The restless mind. *Psychological Bulletin*. 132, 946 (2006).
14. Smallwood, J., Fishman, D., Schooler, J.: Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review*. 14, 230–236 (2007).
15. Graesser, A.: Constructing inferences during narrative text comprehension. *Psychological Review*. 101, 371–95 (1994).
16. McNamara, D., Kintsch, W.: Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*. 22, 247–288 (1996).
17. Ozuru, Y., Dempsey, K., McNamara, D.: Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*. 19, 228–242 (2009).

18. Zwaan, R., Radvansky, G.: Situation models in language comprehension and memory. *Psychological Bulletin*. 123, 162 (1998).
19. Schooler, J., Smallwood, J., Christoff, K., Handy, T., Reichle, E., Sayette, M.: Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*. 15, 319–326 (2011).
20. Smallwood, J., McSpadden, M., Schooler, J.: The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*. 14, 527–533 (2007).
21. Feng, S., D'Mello, S., Graesser, A.: Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin and Review*. (in press).
22. Pekrun, R.: The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*. 18, 315–341 (2006).
23. Deci, E., Ryan, R.: *Intrinsic motivation and self-determination in human behavior*. Plenum: New York and London (1985).
24. Hidi, S., Renninger, K.: The four-phase model of interest development. *Educational Psychologist*. 41, 111–127 (2006).
25. Cordova, D., Lepper, M.: Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*. 88, 715 (1996).
26. Calvert, S., Strong, B., Gallagher, L.: Control as an engagement feature for young children's attention to and learning of computer content. *The American Behavioral Scientist*. 48, 578 (2005).
27. Ainley, M., Hidi, S., Berndorff, D.: Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology*. 94, 545 (2002).
28. Millis, K., Forsyth, C., Butler, H., Wallace, P., Graesser, A., Halpern, D.: *Operation ARIES!: A serious game for teaching scientific inquiry*. Serious Games and Edutainment Applications. Springer, London. 169–195 (2011).
29. Smilek, D., Carriere, J., Cheyne, J.: Out of mind, out of sight Eye blinking as indicator and embodiment of mind wandering. *Psychological Science*. 21, 786–789 (2010).
30. Buhrmester, M., Kwang, T., Gosling, S.: Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*. 6, 3–5 (2011).
31. Mason, W., Suri, S.: Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*. 44, 1–23 (2012).
32. Graesser, A., McNamara, D., Kulikowich, J.: Coh-Matrix providing multilevel analyses of text characteristics. *Educational Researcher*. 40, 223–234 (2011).
33. McNamara, D., Louwerse, M., McCarthy, P., Graesser, A.: Coh-Matrix: Capturing linguistic features of cohesion. *Discourse Processes*. 47, 292–330 (2010).
34. Graesser, A., & Person, N.: Question asking during tutoring. *American Educational Research Journal*, 31, 104–137 (1994).