To read or not to read: Predicting student engagement in interactive reading

Beata Beigman Klebanov, Jonathan Weeks and Sandip Sinharay

ETS, Princeton, USA
bbeigmanklebanov,jweeks,ssinharay@ets.org

Abstract. In this study, upper-elementary-age students used an interactive reading app to read from a classic children's novel during a summer program. Students took turns reading with an adult virtual narrator (audiobook). We use process and background data to explore factors that could predict whether a reader will read their next turn or skip it. We find that skipping quickly becomes self-perpetuating, underscoring the need to support the teacher in providing just-in-time personalized intervention to help students avoid the disengagement trap.

Keywords: children's reading, summer reading, engagement

1 Introduction

The 2022 NAEP results show that 37% of U.S. fourth graders read below the Basic level, an increase since 2019 (34%).[1] There is thus an urgent need to help students recover from the learning loss induced by the pandemic, including a recognition that school-based learning might need to be supplemented after-school and in the summer.[2] Summer enrichment programs can supplement instruction in a more relaxed atmosphere where literacy or math activities are mixed with field trips, games, and other fun activities. However, due to the relaxation of the school discipline and expectations, sustained engagement with learning activities could be a challenge. We analyze data from a summer program where students engaged in independent reading of a children's novel using an electronic shared-reading platform. The research question that we focus on is: What process and demographic factors can help predict reader disengagement?

One of the major pedagogical challenges is classroom orchestration, which refers to design and implementation of classroom activities by the teacher to optimize outcomes. Reviewing the main aspects of orchestration in technology-enhanced learning, [17] pointed out the importance of (a) planning and design, (b) management of the time and workflow, (c) awareness of what is happening in the classroom as a whole and with individual students, (d) adaptation to the emergent occurrences during the learning activity, and (e) interplay of learner-

[1] https://nces.ed.gov/nationsreportcard/reading/
[2] https://njtutoringcorps.org/budget-statement/

and teacher-driven elements of orchestration. Specific students becoming disengaged may be an instance of an emergent occurrence requiring teacher action, as that student's learning may become undermined and the attention of the disengaged student could turn to activities disruptive to others in the classroom.

To help forestall student disengagement, we explore the engagement dynamics in order to understand what characteristics of a student's ongoing and prior activity in the reading app could predict disengagement. If such prediction is possible, it opens up a possibility of supporting effective classroom orchestration by alerting the teacher in real time to allow for a well-timed adaptation action.

2 Related Work

Reader disengagement from a reading activity is often described as 'mind wandering', or following some internal train of thought instead of paying attention to the text; whether or not mind wandering is necessarily unintentional is subject to debate [6]. It is a common phenomenon and has been observed in both reading and listening to a narration of a text, with the second engendering more mind-wandering [10]. While the exact phenomenon in question and methods for detecting it vary substantially [27,21], it is a robust finding in the literature that mind wandering often leads to compromised comprehension of the text [6]: Missing important information early on may lead to failure to build an adequate mental model of the text [23] and cascade further into failure to make relevant inferences later in the reading process [22]. This process may result in a 'vicious cycle', where an initial failure of attention results in impaired comprehension which in turn promotes further inattention [6]. This cyclical view may also suggest some dynamism, where the initial onset of mind wandering might be unintentional but its further maintenance, upon realization of incomprehension, may be a conscious decision.

Both strong and struggling readers may exhibit mind wondering. For the strong readers, mental resources may not be fully engaged with a relatively easy reading activity and start 'working' on something else in parallel. For the struggling readers, the difficulty of the reading may result in the reader disconnecting attention from the reading. In addition to task difficulty, fatigue and lack of interest in the topic have been linked to mind wandering during reading [6]. Individual characteristics such as larger working memory capacity and stronger ability to execute attention control were linked to less mind wandering [25].

Researchers of technology-supported reading investigated detection of mind wandering during literacy activities through monitoring of reading speed [8], eye tracking [7], tracking of scrolling behavior [3], etc. The large bulk of the research is focused on detecting mind wandering as it unfolds or right after it has occurred; however, recent studies also started looking at predicting mind wandering using physiological signs, finding that certain types of arm movements tend to occur about 5 minutes ahead of mind wandering episodes [24]. Prediction of possible mind wandering may help forestall its occurrence, by, for example, helping the reader take a timely break or changing the activity.

While readers can sometimes apply self-initiated recovery strategies, such as re-reading a passage when they realized their minds were wandering, this does not always happen [26]. At the moment, interventions to mitigate mind wandering are generally either reactive (following a detection of mind wandering that is either unfolding or has already occurred) or proactive, focused on techniques such as practicing mindfulness and/or physical activity breaks [14,15], that are most effective when practiced regularly, without a direct relation to any measurements during the reading activity. In a reactive intervention, when eye-tracking detects mind wandering, the reading activity is stopped and a comprehension question is shown; if the response is incorrect, the correct answer is shown to help re-build comprehension; the reader is also prompted to re-read the text [7]. In a different study, eye-tracking-based detection of mind wandering resulted in the reader being prompted for a written self-explanation related to the text content, in order to promote deeper engagement with the content [13].

The psychology literature on mind wandering is generally more concerned with the onset of mind-wandering than with its maintenance (under what conditions it persists [19]), and focuses on brief episodes of disconnect – 10 to 15 seconds. To the best of our knowledge, little is known about the tendency of such episodes to recur, or to develop into a more sustained disengagement. Finally, the reviewed studies were laboratory studies; experts have called for research that would translate laboratory findings into the real world, commenting that it is "an endeavour fraught with complexity and risk, but it is an essential step for research to remain relevant and to contribute to broad societal good" [6].

In this study, we extend the prior research in three ways. First, we investigate a **real-life context** where children read during summer camp, with the attendant lack of control over the experimental environment, including acoustic, technical, and behavioral 'noise', such as construction outside the window, occasionally poor WiFi signal, and anticipation of a subsequent sports or game camp activity, respectively. In such contexts, it may not yet be realistic to use sensitive and expensive equipment with advanced capabilities, although the technology is advancing rapidly [9]. Second, the goal of the reading in our context is not strictly learning and comprehension to a high standard of coherence [4] and the text is not in the expository or informational genres typically used in mind wandering studies [6]. Since this is a relatively **informal reading activity** outside of the regular school context, the target standards of coherence are generally lower and more similar to reading for recreational and entertainment purposes. Third, we investigate a **long-term** engagement with the reading activity, across multiple reading episodes per day and across multiple days. Thus, while we are able to measure aspects of reader activity at a resolution only somewhat inferior to that of the lab studies reviewed above – minutes rather than seconds — we can investigate a much longer-term pattern of engagement.

3 The Interactive Reading App

We built Relay Reader™ [12], a reading and listening app[3], to help developing readers improve fluency while enjoying a good story. Figure 1 shows a screenshot.

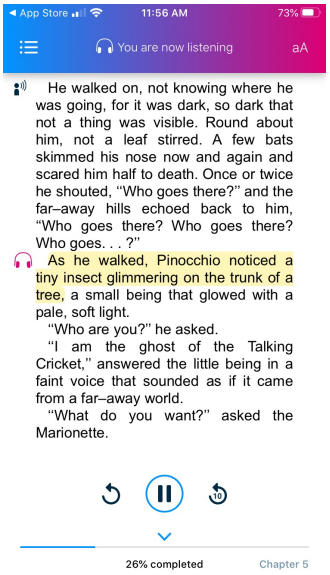

Fig. 1. A screenshot of the Relay Reader app.

The user takes turns reading aloud with a pre-recorded narrator. A turn can be set by the user to between 70 and 200 words on average, separately for the narrator and the user. While the narrator is reading, the text is highlighted for the reader to follow along. The unit of highlight is a **span** – a phrase or a short sentence read by the narrator on one breath. Spans are detected semi-automatically [12]. For example, Figure 1 shows a highlighted 15-word span.

When it is the user's turn to read, the user clicks on a button to start the audio recording and on another button to indicate that they are done. The audio recording is sent for processing to measure performance metrics such as reading accuracy and fluency [2,11]; the measurements are reported to the teacher in a periodic report but are currently not communicated to the user of the app.

After every four turns (two narrator's and two reader's), the reader is asked two multiple-choice **comprehension questions**. The questions are plot-oriented and ask about characters, their appearance, feelings, relationships; about locations and events, or other aspects of the plot. The questions were created manually to ensure effective coverage of important plot elements. The questions are not

[3] https://relayreader.org

meant to trigger deep inference – they probe attention as much as comprehension, and serve a dual purpose of continuously assessing the reader and helping the reader recall an important story element though instant feedback showing the correct answer. In the context of mind wandering research, the app incorporates frequent changes of activity (listening, reading, answering questions) and questions as a mechanism to recover some information that may have been lost if a turn was not read fully or not listened to with attention.

The questions usually ask about information that was explicitly stated (or paraphrased) in the most recent user or narrator turn. We call the latest span in the text that states a piece of information necessary to answer the question the **anchor** of the question; once the bookmark passes the anchor, the question can be asked. After every four turns, the two questions with the nearest preceding anchors are asked. To achieve this local nature of the questions, we created a question for about every 100 words of running text. For example, a question that could be anchored to "I am the ghost of the Talking Cricket" is "Who did Pinocchio meet in the forest?". Note that the anchor does not contain all the information, as it says nothing about the forest, but that information was stated before and so the reader would have been exposed to it prior to this point.

Apart from the recordings of the user's oral reading, the app logs time-stamped events that follow the reader's activity, such as beginning and end of each narrator and student turns, the content of the turn that was read (in narrator's case) or was supposed to be read (in the user's case), responses to the questions. To alleviate any privacy and ethical concerns regarding data collection and usage, we followed our institutions's privacy policy[4] and IRB guidelines.

4 Data

4.1 Data collection context and participants

The data collection occurred in summer programs in 5 locations belonging to the same national organization in the North-East of USA. The 133 participating students were mostly 3-5 graders, split into multi-age groups of 10 per instructor, on average (range: 7-14 students per instructor). While all sites received the same guidelines of 20 minutes of reading 3 times a week, different sites, instructors, and weeks varied – some sessions were longer than others and fewer sessions occurred during some weeks at some sites due to other summer activities. Of the 133 students, 65 were male, 60 female; gender information was not available for 8 students. In terms of grade, 47% of the students finished 4th grade, 27% finished 3rd grade, 17% finished 5th grade, and 5 students (3%) finished 2nd (4) and 1st (1) grades. Grade information was not available for 7 students. All students read *The Adventures of Pinocchio*, a 130-page (39K-word) classic novel by Carlo Collodi (translated from the Italian by Carol Della Chiesa), and listened to the narration by Mark Smith obtained from LibriVox[5] during narrator turns.

[4] https://www.ets.org//legal/privacy.html
[5] https://librivox.org/the-adventures-of-pinocchio-by-carlo-collodi/

4.2 Definitions

The following types of data derived from the logs will be used to compute the variables for our models. Descriptive statistics are shown in Table 1.

Span A phrase or short sentence that can be read on one breath; see section 3 for more details. A span averages 8 words, with a standard deviation of 4.4.
Turn The passage that a participant (narrator or student) is to read out loud.
Turn Length The number of spans in a given turn.
Turn Duration The number of seconds that elapsed while the narrator was reading (for narrator turns); the number of seconds that elapsed between the timestamp of "Start Reading" button pressed by the student and "Done Reading" button pressed by the student (for student turns).
RCQ A comprehension question asked during the activity; see section 3.
Skipped Turn If the duration of a turn is such that it would have taken less than one second per span had the reader actually read it, the turn is extremely unlikely to have been fully read, as it would entail a reading rate of about 480 words per minute (60 sec $\times$ 8 words) – an unrealistic reading rate, even for silent reading.[6] We consider such turns skipped. Note that it is not possible to skip a narrator's turn, only student turns can be skipped.

Table 1. Descriptive statistics of the dataset, per reader, after the pre-processing described in section 5.2. The data come from $n = 133$ readers.

Variable	Mean	Stdev	Min	Max
# Reading turns	54.6	27.5	2	104
# Reading turns per day	6.5	4.3	2	20
# Skipped reading turns	18.6	18.8	0	83
# Spans per turn	15.9	5.5	0	40
RCQ % correct	65.3	22.6	0	100
Duration of a reading turn (in seconds)	62.3	49.5	2.9	281.1

5 Predicting Disengagement

5.1 Independent variables (predictors)

Our research question relates to finding process and demographic factors that may predict disengagement. We operationalize the prediction problem as one of estimating the probability of a student skipping their upcoming reading turn given the characteristics of the immediate task, of the student's reading behavior so far and demographic variables. The problems that process variables are trying to capture are described below. Table 2 shows turn-level statistics.

[6] Most adults read English fiction silently at a rate of 200-320 words per minute [5].

1. *I am already disengaged*: A variable that counts the number of skips in the three most recent reading turns: **SkipsIn3**.
2. *I am getting fatigued by this activity*: The number of turns already read today is counted by the **UnitDay** variable; every two turns form a unit.
3. *I am doing too big a share of the reading*: The ratio of the length of the most recent narrator turn to the length of the upcoming student's turn: **NSRatio**.
4. *I am not quite following the story*: Proportion of RCQs answered incorrectly so far: **PIncorrect**.

Note that UnitDay, NSRatio, and PIncorrect correspond to variables found to be related to mind wandering in the literature – fatigue, an aspect of task difficulty (in a student's perception), and reading comprehension. The SkipsIn3 variable addresses the dynamic of the reading engagement in time and has not been considered in mind wandering research so far, to the best of our knowledge.

Table 2. Descriptive statistics per user reading turn; n = 7,208 turns.

Variable	Mean	Stdev	Min	Max
SkipsIn	1.03	1.27	0	3
NSRatio	1.80	1.52	0.33	21
UnitDay	3.23	2.15	1	10
PIncorrect	0.34	0.23	0	1

We also consider grade and gender. If grade has predictive value, for example, if it is harder for younger students to sustain attention, instructors in the multi-grade classrooms might stop the activity earlier for younger readers. Some prior research and public discourse suggest that boys may have more difficulty engaging with reading, especially of fiction; see [20] for a critique.

5.2 Data pre-processing for analysis

The data were structured using a person-period format where all data associated with a particular student reading turn appear in a row. That is, the number of rows per student corresponds to the number of student reading turns. Indicator variables were included for day and unit within day. Lagged variables for three previous turns were merged to the records for each current turn. For instance, the row with data for turn 4 includes the data for turns 3, 2, and 1. We also created cumulative variables; their values are accumulated up to the given time point. Lastly, since students seldom read more than 20 passages in a day, the data were truncated to only include up to the 10th UnitDay (20th reading turn).

5.3 Models

The outcome of interest is the odds of a student skipping during the upcoming turn. Given that there are multiple observations for each student, it is important

to take into account within-student clustering. As a first step, we fit a two-level mixed-effects model with random intercepts and no additional predictors. The intraclass correlation was 0.7, which suggests that clustering at the student level will lead to more accurate statistical inference. As a next step, we examined whether there was meaningful variation in the effect of time. We considered random effects for observations nested within days and observations nested within time points. The former relates to potential clustering over shorter time intervals, whereas the latter – over the full duration of participation. In both cases, the intraclass correlations were very small (around 0.015), suggesting that the probability of skipping is not expected to vary much from day to day. We therefore opted for a two-level model with fixed effects for various predictors and random intercepts for the students.

In the subsequent steps, we examined several models. For each model, we considered the significance of the predictors, the practical significance of the coefficients, the increase in pseudo R^2 [16], the impact on model fit (reduction in AIC and BIC), and the correlation between the fixed effects (an indicator of multicollinearity).

We fitted generalized linear mixed models to the data. The models were implemented using a logit link function; the parameters were estimated via maximum likelihood (Laplace approximation) with the `lme4` [1] package in R [18]. The dependent variable is a binary indicator of whether the current reader's turn is skipped. The odds of student i skipping are:

$$odds = P(Skip_i = 1)/(1 - P(Skip_i = 1)) \tag{1}$$

The baseline model was specified as in Equation 2, where β_{00} is the average log-odds and u_{0i} is the deviation for the cluster-specific log-odds:

$$logit(odds) = \beta_{00} + u_{0i} \tag{2}$$

In the next model we specified fixed effects for the six variables listed in section 5.1 and a random effect for the intercept with `student` at level-2. In this model, all of the process variables were significant predictors of skipping while neither of the demographic variables was a significant predictor. We therefore removed the demographic variables from the model and refitted the following final model specified in Equation 3:

$$\begin{aligned} logit(odds) = \beta_{00} + \beta_{10} SkipsIn3_i + \beta_{20} UnitDay_i + \\ + \beta_{30} NSRatio_i + \beta_{40} PIncorrect_i + u_{0i} \end{aligned} \tag{3}$$

6 Results & Discussion

The baseline model with the random effect only (Equation 2) fits with AIC = 5,486, BIC = 5,500, and Log Likelihood = -2,741. Table 3 shows summary statistics for the model in Equation 3. The log likelihood of the model is larger

Table 3. Model estimates and fit statistics for the model specified in Equation 3 for predicting the probability of a reader skipping their current reading turn. n = 7,208.

	Variance	St. Dev.	
Random Effects			
Student	1.92	1.387	
	Coefficient	Signif.	St. Error
Fixed Effects			
Intercept	-2.98	$p < 0.001$	(0.192)
SkipsIn3	1.01	$p < 0.001$	(0.041)
UnitDay	0.16	$p < 0.001$	(0.020)
NSRatio	-0.27	$p < 0.001$	(0.030)
PIncorrect	1.69	$p < 0.001$	(0.260)
Fit Statistics			
AIC		4,528	
BIC		4,570	
Log Likelihood		-2,258	
Pseudo $R^2(Fixed + Random)$		0.82	
Pseudo $R^2(Fixed)$		0.49	

than that of the baseline model by 18%. The fixed effects explain 49% of the variance in the observations, as measured by pseudo R^2 for the fixed effects [16].

The model coefficients for each fixed effect show the odds ratio of skipping to not-skipping the current turn controlling for all the other variables. For example, each unit read within a day adds 0.16 to the skipping odds; thus, all else being equal, skips are likelier later in the daily reading session. If the narrator's latest turn is of about the same length as the student's (NSRatio close to 1), the odds of skipping are reduced by 0.27. Skipping each of the preceding three turns increases the odds of skipping the next one by 1. Having low story comprehension increases the likelihood of skipping (note that the variable codes for percent *in*correct).

We observe that the signs of the coefficients of the variables that implement literature-based constructs align with results reported in prior work – low comprehenders and students who are becoming fatigued from the activity are likelier to skip; if the upcoming turn is substantially shorter than the preceding narrator's turn, which may be perceived by the student as having a much easier job to do than the narrator, the reader is less likely to skip the turn. These results provide evidence for generalization of the findings in the literature to a real-life extended reading context in an informal educational settings.

Our most sobering finding is that disengagement is something that a student can get trapped in very quickly. Table 4 shows the probabilities of skipping given different number of skips within the last three turns. A reader who has not skipped any of the preceding three turns is highly unlikely to skip the next turn. In contrast, having skipped the three immediately preceding turns, a reader with average comprehension (65%) has a 40% chance of skipping the next turn – even if fatigue is low and the reading is balanced between reader and narrator.

Table 4. Predicted probability of skipping the current turn for a 65%-comprehension reader as a function of skipping none, one, two, or three of the three preceding turns and of the unit in the day. X means the model predicts no skipping (negative odds).

	Probability of skipping			
Unit	***for an average-comprehension reader***			
Day	SkipIn3=0	SkipIn3=1	SkipIn3=2	SkipIn3=3
2	X	X	X	0.40
3	X	X	X	0.45
4	X	X	X	0.49
5	X	X	0.11	0.53
6	X	X	0.22	0.56
7	X	X	0.30	0.59
8	X	X	0.37	0.61
9	X	X	0.43	0.64
10	X	X	0.47	0.66

The model is not successful in predicting a reader's first skip. Without prior skips, all reasonable constellations of other variables predict no skipping (due to the strong negative intercept).[7] Given that there is no useful predictor of the first skip and given the high odds of further skipping following three consecutive skips, our findings underscore the importance of a good design of the activity based on the model and of real-time adaptation of the activity in a personalized manner if some students do show a snowballing disengagement pattern.

To exemplify the way the model informs activity design, let us consider Table 5 that shows the importance of the fatigue and story comprehension factors. Two skips early in the session might not set the reader up for subsequent skipping, but the chances of skipping increase substantially later in the session for readers with low and average comprehension. Recall that one unit corresponds to two reading turns; due to the interleaved nature of the activity, for there to be two reading turns, there also must have been two narrator turns. Assuming a 150-word average narrator turn and 100-word average student turn, these correspond, roughly, to the amount of text that could be read in about one minute (According to Table 1, average student turn lasted 62.3 seconds). Thus, a unit of the activity is expected to take about 4 minutes of net reading and listening, plus short breaks, responses to questions, re-plays if necessary – about 5 minutes of activity. According to Table 5, on the 6th unit of the day, that is, after about 25 minutes of the activity, students who skipped two preceding turns and have low story comprehension have 38% chance of skipping the next turn. Perhaps 20-25 minutes would be a good target duration of this activity for the class. If the scheduling can be personalized, students with stronger story comprehension could continue for a few more turns.

[7] The intercept of -2.98 is not compensated even if it is the 10th unit of the day (10*0.16 = 1.6) and the reader is guessing RCQs (PIncorrect=0.75, the odds increase is 1.27), 1.27+1.60 = 2.87 < |-2.98|, so the model would still predict no skipping.

Table 5. Impact of fatigue (UnitDay) and story comprehension (RCQ percent correct) on predicted probabilities of skipping after skipping two of the preceding three turns.

Unit	***Probability of skipping the next turn given 2 skips***		
Day	45% RCQ	65% RCQ	85% RCQ
2	X	X	X
3	0.13	X	X
4	0.23	X	X
5	0.32	0.11	X
6	0.38	0.22	X
7	0.44	0.30	0.09
8	0.48	0.37	0.20
9	0.52	0.43	0.29
10	0.55	0.47	0.36

To exemplify opportunities for a personalized adjustment of the activity, Table 6 illustrates the impact of the variable that captures the ratio of the length of the narrator's turn to the student's. According to the model, a high ratio predicts lower probability of skipping. For example, a student with low story comprehension has a 23% chance of skipping on unit 4 if the narrator and the student have turns of approximately equal length, which would go down to 3% if the next student turn were only about half as long as the narrator's (see columns 2,3 row 2 in Table 6). Likewise, a student with average story comprehension who skipped two turns would have a 22% chance of skipping the next turn in the equal turn length situation on unit 5, but the probability would go down to only 1% if the next reading turn were much shorter than that narrator's turn (see columns 4,5 row 3 in the Table). Since narrator and student turn length are separately adjustable in the app, a shorter reader turn and a longer narrator turn can be arranged to help keep the student on track. It is also possible that other teacher actions, such as asking the student to pause and conversing with them, would provide a break in the activity and help redirect the trajectory of subsequent reading away from further skipping.

Table 6. Illustration of the impact of the ratio of the narrator to student turn length (NSRatio). N = S corresponds to a case where narrator and student turns are approximately of the same length; in N = 2S the narrator is reading twice the amount.

	Probability of skipping the next turn given 2 skips			
Unit	45% RCQ		65% RCQ	
Day	N = S	N = 2S	N = S	N = 2S
2	X	X	X	X
3	0.13	X	X	X
4	0.23	0.03	0.11	X
5	0.32	0.16	0.22	0.01

Finally, we observe that the background variables we explored – grade and gender – were not significant predictors of skipping. This result suggests that all the students in the grade range we explored – mostly 3-5 graders – were equally able to engage with the activity, which may make it easier to administer in a multi-age group of readers. We also found no evidence that boys are less likely to be engaged in reading fiction – contrary to a common belief.

7 Conclusion

In this study, upper-elementary-age students used an interactive reading app to read from a children's novel during a literacy enrichment part of a summer program. Students took turns reading with an adult virtual narrator (audiobook). We used process data to explore factors that could predict whether a reader will read their upcoming reading turn or skip it. We found that prior skipping is the strongest predictor of future skipping, suggesting that disengagement is likely to snowball if left unattended.

We illustrated a way the model can be used to estimate a reasonable duration of the activity and to plan for a personalized activity adjustment action in case continued disengagement hazard is flagged for a particular student. Our results point towards the existence of a window of opportunity for a teacher's corrective action – between the first observed skip and the next one or two, since the first cannot be predicted and the third is predictive of a high likelihood of a reader getting trapped in a sustained pattern of disengagement. After a flagged first skip, the teacher may have only about 5 minutes to act before the student gets to the third consecutive skip. Identifying possible effective teacher actions for various types of students is an urgent area for future research.

Our results suggest that factors reported in the literature as conducive to mind wandering during reading, namely, fatigue, task difficulty, and low comprehension, generalize to the new context examined in this study – a relatively informal reading of a novel by upper elementary students during summer camp. Our findings also show a quick transition from a one-off to sustained disengagement, thus providing empirical evidence to the 'vicious cycle' hypothesis discussed in the mind wandering literature [6]. Collecting further empirical evidence, with readers of different ages reading in a different context, is also an important avenue for future work in order to deepen the understanding of the fine-grained dynamics of the process of reader disengagement, which, in turn, could lead to improved prevention, detection, and recovery solutions.

8 Limitations

A number of limitations of the current work are recognized; they all point to directions for future work to examine the robustness of the findings. One limitation is the use of one book; other reading materials may generate different engagement patterns. Secondly, the activity took place in a multi-age group; a more agewise homogeneous group could show a different engagement pattern.

9 Acknowledgment

This material is based upon work partially supported by the National Science Foundation and the Institute of Education Sciences under Grant #2229612. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the U.S. Department of Education.

This version of the contribution has been accepted for publication in the proceedings of AIED 2024 after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/978-3-031-64299-9_15. Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use https://www. springernature.com/gp/open-science/policies/accepted-manuscript-terms.

References

1. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. Journal of Statistical Software **67**(1), 1–48 (2015)
2. Beigman Klebanov, B., Loukina, A.: Exploiting structured error to improve automated scoring of oral reading fluency. In: Proceedings of AIED. pp. 76–81 (2021)
3. Biedermann, D., Schneider, J., Ciordas-Hertel, G.P., Eichmann, B., Hahnel, C., Goldhammer, F., Drachsler, H.: Detecting the disengaged reader-using scrolling data to predict disengagement during reading. In: LAK23: 13th International Learning Analytics and Knowledge Conference. pp. 585–591 (2023)
4. Van den Broek, P., Lorch, R., Linderholm, T., Gustafson, M.: The effects of readers' goals on inference generation and memory for texts. Memory & cognition **29**(8), 1081–1087 (2001)
5. Brysbaert, M.: How many words do we read per minute? a review and meta-analysis of reading rate. Journal of Memory and Language **109**(104047) (2019)
6. D'Mello, S., Mills, C.: Mind wandering during reading: An interdisciplinary and integrative review of psychological, computing, and intervention research and theory. Language and Linguistics Compass **15**(4), e12412 (2021)
7. Faber, M., Bixler, R., D'Mello, S.: An automated behavioral measure of mind wandering during computerized reading. Behavior Research Methods **50**, 134–150 (2018)
8. Franklin, M.S., Smallwood, J., Schooler, J.W.: Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. Psychonomic bulletin & review **18**, 992–997 (2011)
9. Hutt, S., Wong, A., Papoutsaki, A., Baker, R., Gold, J., Mills, C.: Webcam-based eye tracking to detect mind wandering and comprehension errors. Behavior Research Methods pp. 1–17 (2023)
10. Kopp, K., D'Mello, S.: The impact of modality on mind wandering during comprehension. Applied Cognitive Psychology **30**(1), 29–40 (2016)
11. Loukina, A., Beigman Klebanov, B., Lange, P.L., Qian, Y., Gyawali, B., Madnani, N., Misra, A., Zechner, K., Wang, Z., Sabatini, J.: Automated estimation of oral reading fluency during summer camp e-book reading with My Turn To Read. In: Proceedings of INTERSPEECH. pp. 21–25 (2019)

12. Madnani, N., Beigman Klebanov, B., Loukina, A., Gyawali, B., Lange, P.L., Sabatini, J., Flor, M.: My Turn to Read: An interleaved e-book reading tool for developing and struggling readers. In: Proceedings of the ACL: System Demonstrations. pp. 141–146 (2019)
13. Mills, C., Gregg, J., Bixler, R., D'Mello, S.: Eye-mind reader: an intelligent reading interface that promotes long-term comprehension by detecting and responding to mind wandering. Human–Computer Interaction **36**(4), 306–332 (2021)
14. Mrazek, M., Zedelius, C., Gross, M., Mrazek, A., Phillips, D., Schooler, J.: Mindfulness in education: Enhancing academic achievement and student well-being by reducing mind-wandering. In: Mindfulness in social psychology, pp. 139–152. Routledge (2017)
15. Müller, C., Otto, B., Sawitzki, V., Kanagalingam, P., Scherer, J.S., Lindberg, S.: Short breaks at school: effects of a physical activity and a mindfulness intervention on children's attention, reading comprehension, and self-esteem. Trends in Neuroscience and Education **25**, 100160 (2021)
16. Nakagawa, S., Schielzeth, H.: A general and simple method for obtaining r2 from generalized linear mixed-effects models. Methods in ecology and evolution **4**(2), 133–142 (2013)
17. Prieto, L., Holenko Dlab, M., Gutiérrez, I., Abdulwahed, M., Balid, W.: Orchestrating technology enhanced learning: a literature review and a conceptual framework. International Journal of Technology Enhanced Learning **3**(6), 583–598 (2011)
18. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2019), https://www.R-project.org
19. Randall, J., Oswald, F., Beier, M.: Mind-wandering, cognition, and performance: a theory-driven meta-analysis of attention regulation. Psychological bulletin **140**(6), 1411 (2014)
20. Scholes, L., Spina, N., Comber, B.: Disrupting the 'boys don't read' discourse: Primary school boys who love reading fiction. British Educational Research Journal **47**(1), 163–180 (2021)
21. 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**(7), 319–326 (2011)
22. Smallwood, J.: Mind-wandering while reading: Attentional decoupling, mindless reading and the cascade model of inattention. Language and Linguistics Compass **5**(2), 63–77 (2011)
23. Smallwood, J., McSpadden, M., Schooler, J.: When attention matters: The curious incident of the wandering mind. Memory & cognition **36**, 1144–1150 (2008)
24. Southwell, R., Peacock, C., D'Mello, S.: Getting the wiggles out: Movement between tasks predicts future mind wandering during learning activities. In: Proceedings of AIED. pp. 489–501 (2023)
25. Unsworth, N., McMillan, B.: Mind wandering and reading comprehension: examining the roles of working memory capacity, interest, motivation, and topic experience. Journal of Experimental Psychology: Learning, Memory, and Cognition **39**(3), 832 (2013)
26. Varao-Sousa, T., Solman, G., Kingstone, A.: Re-reading after mind wandering. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale **71**(3), 203 (2017)
27. Weinstein, Y., De Lima, H., Van Der Zee, T.: Are you mind-wandering, or is your mind on task? the effect of probe framing on mind-wandering reports. Psychonomic Bulletin & Review **25**, 754–760 (2018)