Check for updates

ets research institute

DECEMBER 2024
RR–24-14

RESEARCH REPORT

Monitoring Oral Reading Fluency From Electronic Shared Book Reading

Insights From a Full-Length Book Reading Study With Relay Reader®

AUTHORS

Zuowei Wang, Beata Beigman Klebanov, Tenaha O'Reilly, and John Sabatini

ETS Research Report Series

EIGNOR EXECUTIVE EDITOR

Daniel F. McCaffrey
Lord Chair in Measurement and Statistics

ASSOCIATE EDITORS

Usama Ali
Senior Measurement Scientist

Beata Beigman Klebanov
Principal Research Scientist, Edusoft

Heather Buzick
Senior Research Scientist

Katherine Castellano
Managing Principal Research Scientist

Larry Davis
Director Research

Paul A. Jewsbury
Senior Measurement Scientist

Jamie Mikeska
Managing Senior Research Scientist

Teresa Ober
Research Scientist

Jonathan Schmidgall
Senior Research Scientist

Jesse Sparks
Managing Senior Research Scientist

Zuowei Wang
Measurement Scientist

Klaus Zechner
Senior Research Scientist

Jiyun Zu
Senior Measurement Scientist

PRODUCTION EDITORS

Kim Fryer
Manager, Editing Services

Ayleen Gontz
Senior Editor/Communication Specialist

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ETS Research Report Series ISSN 2330-8516

ETS RESEARCH REPORT

Monitoring Oral Reading Fluency From Electronic Shared Book Reading: Insights From a Full-Length Book Reading Study With Relay Reader®

Zuowei Wang[1], Beata Beigman Klebanov[1], Tenaha O'Reilly[1], & John Sabatini[2]

1 ETS Research Institute, ETS, Princeton, New Jersey United States
2 Institute for Intelligent Systems, University of Memphis, Memphis, Tennessee United States

Existing research reveals a robust relationship between self-reported print exposure and long-term literacy development, yet few studies have demonstrated how reading skills change as children read a book in the short term. In this study, 50 children (mean age 9.7 years, *SD* = .8) took turns with a prerecorded narrator reading aloud a popular children's novel, producing 6,092 oral reading responses over 1,093 book passages. Each oral reading response was evaluated by a speech engine that calculated words-correct-per-minute (WCPM). Mixed effect models revealed that text level differences, between-individual differences, and within-individual variations explained 13%, 56% and 32% of variance in WCPM, respectively. On average, children started reading the book at about 93 WCPM, and they improved by 2.26 WCPM for every 10,000 words of book reading. Random effects showed that the standard deviation of the growth rate was 1.85 WCPM, suggesting substantial individual difference in growth rate. Implications for reading instruction and assessment were discussed.

Keywords Book reading; oral reading fluency; reading development; mixed effects model; Relay Reader®

doi:10.1002/ets2.12390

It is widely believed that reading books benefits literacy development (Anderson et al., 1988; Baker et al., 1997; Bus et al., 1995; Hargrave & Sénéchal, 2000; McBride-Chang et al., 1993; Scarborough & Dobrich, 1994; Torppa et al., 2020). Indeed, research has indicated that self-reported book reading is associated with positive reading comprehension and language outcomes (Bus et al., 1995; Scarborough & Dobrich, 1994; Torppa et al., 2020). In a meta-analysis, Mol and Bus (2011) found that print exposure was moderately related to reading comprehension ($r = .36$) based on 21 studies focusing on students from Grades 1 through 12. However, research studies that implemented interventions encouraging more book reading have had more limited success (Ardoin et al., 2016; Carver & Leibert, 1995; Ingham, 1982; Kim & Guryan, 2010). With these mixed findings, some have expressed doubts about the efficacy of encouraging children to read more books to develop their reading skills (Baker et al., 1998; Carver & Leibert, 1995; Harlaar et al., 2011; Scarborough & Dobrich, 1994; What Works Clearinghouse et al., 2015).

Collectively, these studies examined the impact of exposure to a wide range of books and texts across a span of months to years on distal measures of reading skills. However, what we do not have are finer grained empirical data on how reading a single book may contribute to reading skill development. Addressing this question requires a more direct, data-intensive approach that more precisely examines what is read, what processes are engaged by the reader, and how reading ability develops in the process. Recent advances in technology such as electronic book reading platforms, speech processing, and natural language processing have made this data-intensive approach possible.

The goal of this study was to explore changes in oral reading fluency (ORF) when elementary school students engage in sustained reading of a book. We invited children to read on Relay Reader®, an electronic shared book reading platform designed to provide an engaging reading experience meanwhile capturing key information about children's reading behavior. Children took turns with a narrator reading aloud the book, and their oral reading responses were collected and processed by a speech processing engine to monitor ORF changes.

Corresponding author: Zuowei Wang, Email: zwang@ets.org

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Individual Differences in Oral Reading Fluency and the Role of Book Reading

ORF has long been used as an indicator of children's reading competence (Fuchs et al., 2001), and it is a strong predictor of students' reading comprehension (Roehrig et al., 2008) especially among primary-grade students (Wayman et al., 2007). In a typical measure of ORF, the student is asked to read aloud a passage, and performance is calculated as the number of words the student reads aloud correctly in a minute, namely, words-correct-per-minute (WCPM). A high WCPM score indicates the reader can orchestrate a number of component skills at ease, including decoding words, accessing lexical information, connecting information within and between sentences to make sense of the text (Fuchs et al., 2001).

Large individual differences in ORF emerge early in the school years. At the beginning of Grade 3, the difference in ORF between students at the 90th percentile and students at the 10th percentile is already over 100 WCPM, and this difference persists to the end of Grade 8; a first grader reading at the 90th percentile has similar WCPM scores to that of a sixth grader at the 10th percentile (Hasbrouck & Tindal, 2006).

The large individual difference in ORF may be partly attributable to the amount of reading practice that students engage in. Indeed, prior research shows that children who spend more than 30 minutes a day reading books show substantially faster reading development than children who spend less than 15 minutes a day reading (Renaissance Learning, 2019), and the amount of book reading students engage in predicted their later reading comprehension development (Torppa et al., 2020). Furthermore, the effect of print exposure on language development increases from kindergarten through high school (Mol & Bus, 2011). As Mol and Bus (2011) wrote, "The outcomes support an upward spiral of causality: Children who are more proficient in comprehension and technical reading and spelling skills read more; because of more print exposure, their comprehension and technical reading and spelling skills improved more with each year of education" (p. 267). They go on to state: "Interestingly, poor readers also appear to benefit from independent leisure time reading" (p. 267).

On the other hand, after reviewing studies of shared book reading that met its standards, the U.S. Department of Education What Works Clearinghouse concluded that based on existing evidence shared book reading had no discernible effects on alphabetics or general reading achievement and mixed effects on comprehension and language development (What Works Clearinghouse et al., 2015). One potential reason for why there was no discernible effect uncovered in these studies may have to do with the types of designs employed. For example, one thing these studies have in common is that they all used a pretest, intervention, posttest design. This type of design has a limited number of observations that might have missed opportunities to detect any reading gains. Support of this interpretation comes from a study by Ardoin et al. (2016), who monitored the impact of wide reading and repeated reading on reading gains by giving participants weekly ORF assessments over the course of about 10 weeks. With many more observations, the results showed that students in both conditions demonstrated significantly more gains in ORF than those in a business-as-usual control group. Inspired by Ardoin et al., this study utilized continuous ORF monitoring to help detect the impact of book reading on children's reading skills.

Leveraging Shared-Book Reading to Promote Engagement and Development

For children who are yet to develop adequate reading skills to enjoy reading a book on their own, shared book reading can be an effective method to keep them engaged with the book content, meanwhile providing the modeling for children to develop their reading skills (Blewitt et al., 2009). Different variations of shared book reading activities have been used to promote reading development across age groups, including joint book reading (Bus et al., 1995), assisted reading (Esteves & Whitten, 2011), and paired reading (Topping, 1987). The level of child involvement differs in these practices. Notably, shared book reading can be performed not only between a child and a care-giver, but also between a child and an audiobook (Esteves & Whitten, 2011). The audiobook narration, often provided by a professional narrator, provides an excellent reading model for the child to learn from.

In this study, we utilized Relay Reader to facilitate shared book reading between the child and audiobook narration. The platform has been shown effective in keeping children engaged with the extended book reading activity (Beigman Klebanov et al., 2019). When reading on Relay Reader, the child alternates between listening to a skilled, expressive narration (i.e., the narrator's turn) to engage their interest in the story and reading passages aloud themselves (i.e., the child's turn). During the narrator's turn, children's attention is directed to the corresponding text with the sentence or phrases being read aloud by the narrator highlighted. During the child's turn, the audiobook narration stops playing and the child

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

is prompted to continue to read aloud. The beginning and the end of the passage that the child is expected to read for their turn is highlighted. After every other reading turn, the child is prompted to answer one or two multiple-choice comprehension questions based on what they or the narrator have just read aloud. Once they have selected an answer, the correct answer is provided. Their accuracy on all comprehension questions they have answered is provided to students as feedback, with the goal of encouraging them to read for understanding.

Such a design enables a balance of demands between the narrator modeling fluent reading and the child practicing reading. During the narrator's turn, the child is expected to engage in assisted reading (Esteves & Whitten, 2011) by listening to the narration and silently reading the text. This provides the child with the opportunity to learn how unfamiliar words are pronounced and spelled (Blewitt et al., 2009). Furthermore, the narrator's turn provides the child with an opportunity to develop a mental model of the book content, even if the child's text reading ability is so effortful that comprehension when attempting to read themselves is compromised. This is especially important to keep the child engaged with the book (Renaissance Learning, 2018). During the child's reading turn, the child is expected to practice reading aloud and hence develop ORF. Relevant to this study, because the child is reading aloud, this provides opportunities to measure and monitor the child's fluency growth.

Monitoring Oral Reading Fluency With Non-Standardized Book Passages

Traditional ORF assessments use standardized probe passages to make the WCPM scores comparable across passages. For example, one of the most extensively researched and widely used ORF assessments, DIBELS (Good et al., 2013), used the DMG Passage Difficulty Index to select passages so that passages within any grade level are comparable in each of the three dimensions: decoding difficulty, semantic difficulty, and syntactic difficulty. DIBELS developers then conducted empirical studies collecting student performance data reading aloud these preselected passages (Powell-Smith et al., 2010). Passages that were read aloud by students with performance that substantially deviated from other passages were eliminated (Good et al., 2013). T his passage selection process helps assemble a set of passages for each grade level that are homogeneous in readability based on measures provided by the DMG Passage Difficulty Index.

In contrast, passages in a book are less similar than one might think and therefore less comparable by design (Beigman Klebanov et al., 2017; Eberwein, 1979; Liceralde et al., 2022). Previous research shows that a single book contains a wide range of text complexity, more than the range of suggested grade levels that is of ten used to describe a book. For instance, in one study the authors found that text sections excerpted from a popular children book, *Harry Potter and the Sorcerer's Stone*, ranged from Grade 2 to Grade 12. Such variation in text features can also contribute to variation in students' WCPM scores, which if left unaccounted for might result in bias when estimating the growth trajectory of students' ORF (Barth et al., 2014; Francis et al., 2008). In the current study, we analyzed text features of book passages and compared the results to text features of probe passages used in traditional ORF assessments (RQ1). Because we found book passages had a wider range of text features than traditional ORF probe passages, we analyzed variance components in students' WCPM scores to examine the proportion of WCPM variances due to passage differences vs. individual differences with a cross-classified random effect model that treated both person and passage as random effects (RQ2).

To help mitigate the impact of nonstandardized passages on the growth estimate of students' WCPM scores, this study adopted a continuous monitoring approach, wherein each student on average provided over 100 oral reading responses. In a simulation study, Christ et al. (2012) demonstrated that a greater number of data points could compensate for large residual variance in WCPM when estimating ORF growth. For example, the reliability of ORF growth estimate on a data set with smaller residual variance in WCPM (i.e., 5 WCPM) and consisting of 15 observations, is similar to that of another data set with larger unexplained variance (i.e., 10 WCPM) and consisting of 20 observations. These simulation results suggest that thanks to the large number of observations, the current study may have a higher tolerance of residual variance in WCPM due to non-standardized passages. To test this hypothesis, we compared the level of unexplained residual variance in WCPM to Christ et al. (2012) when estimating the growth rate in students' WCPM (RQ3).

Summary and Research Questions

The relationship between print exposure and literacy development is well documented in long-term correlational studies that often use, less than desirable, self-report measures (Mol & Bus, 2011). However, when educators investigate the specific impact of reading an actual book on reading development within a shorter time frame, this relationship becomes

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

much less clear. We believe a more data intensive approach to detecting growth will help clarify the inconsistent results. Created and used as a progress monitoring tool, ORF holds promise for such a measure (Christ et al., 2012; Deno, 1985). In this study, we explored whether we could detect changes in Grade 3–5 students' ORF, as reflected by the WCPM on passages that they read aloud in the process of completing reading a full-length novel on Relay Reader. We address the following research questions (RQs):

1. What is the range of readability of passages in the book, and how does it compare to the range of readability in standardized ORF assessments?
2. Given the wide range of readability of book passages, how much variation in students' WCPM is contributed by text differences vs. individual differences?
3. Do students' WCPM improve as they read the book?

Method

Participants

Fifty students completed a book reading activity, during which they provided 20 or more scorable oral reading responses. These students came from a private elementary school located in a northeastern U.S. state. The school used Relay Reader as part of their daily activities. We did not select the study site for specific reasons other than that the school principal agreed to use a pilot version of Relay Reader. Students were allowed to participate if their instructors or parents determined the reading activity would be appropriate for them and guardian consent was provided to us. We did not otherwise pose any inclusion/exclusion criteria. Since the book reading activity was designed to be self-administered, we did not provide extra instructions to the school for the book reading activity.

Demographic information was available from 49 of the 50 students: their mean age was 9.7 years old ($SD = 0.7$) at the beginning of the book reading activity; 26 students were female, and 23 were male. Eight students were from Grade 3, 26 were from Grade 4, and 15 were from Grade 5. The study was reviewed and approved by the Committee for Prior Review of Research at ETS. Prior to participating, legal guardians' consent was obtained from each student.

Materials and Procedures

Book Reading Activity

Students read the book *Harry Potter and the Sorcerer's Stone* (HP; Rowling, 1999) with an earlier version of Relay Reader under a research license granted by the publisher. This book was selected due to its popularity among the age range, which we hoped would facilitate student engagement when participating in the reading activity. According to Scholastic, the U.S. publisher of this book, it is appropriate for children in Grades 3–8.

Using Relay Reader, the child alternates between listening to the narrator (narrator's turn) and reading aloud (child's turn). During a narrator's turn, the child listens to a passage read aloud by the narrator and can also read along. During a child's turn, the child is expected to read aloud a passage assigned to them. The boundary of passages across turns is predetermined based on the user's setting of the length of the child's reading turns. Additionally, each chapter always starts from the narrator's turn, and each time the child logs on to Relay Reader, the child always starts from the narrator's turn: Because of these implementations, as children read more at different paces, they diverged in the exact set of passages they listened to or read aloud.

To encourage reading for understanding, the child is prompted to answer one or two multiple-choice comprehension questions based on what the child or the narrator has just read aloud. Once the child has provided an answer, the correct answer is provided as feedback. All questions are designed to be simple, focusing on readers' literal understanding. The answers to these questions can be found from a single sentence in the passages the child has just read aloud or listened to.

When the child reads aloud a passage in their reading turn, their oral reading response is recorded on Relay Reader and transferred to a server for speech processing, which yields a WCPM score, thus providing one observation of the child's ORF.

In this study, students were asked to read the book on Relay Reader for about 20 minutes a day, three to four days per week, and the reading activity extended up to 19 weeks, depending on when the student finished reading

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the book. The average time for this sample to finish reading the book was about 10 weeks. Most student turns were fixed to about 150 words, with shorter or longer turns possible to preserve the natural boundary of paragraphs and chapters. The book reading activity was group administered. Each student wore a headphone with a microphone. Students read at their own pace without interacting with others during book reading. Students were encouraged by their teachers to read aloud and to read for understanding. In each reading session, students read the book on Relay Reader in a low stakes natural setting until time was up for the session, when they would move on to other activities.

Text Features of Book Passages

In each child's reading turn, the child is assigned to read aloud a passage. Their oral reading of the passage is an oral reading response that provides one observation of their ORF. A total of 1,093 unique HP passages that were assigned as child's reading turns (as opposed to narrator turns) on Relay Reader had scorable oral reading responses from students. Some of these passages had overlap with other passages. We processed these passages using TextEvaluator® (Sheehan et al., 2014), which provided the grade level estimate of each passage. TextEvaluator predicts a passage's grade level by applying regression models that use text features along eight dimensions as predictors. For more details, please see Sheehan et al. (2014).

Automatic Scoring of Oral Reading Responses

Students' oral reading responses in Relay Reader were processed using an automatic scoring engine for ORF that has been validated for this context. The engine transcribes an oral reading response and then matches the transcribed text to the corresponding book passage to calculate the number of words that are correctly read aloud. This number is then divided by the duration of the transcribed oral reading response, resulting in WCPM. Prior research demonstrated that the automatic scoring engine provided satisfactory performance, with automatic WCPM scores correlating to manual WCPM scores at $r = .93$ at the individual student level, who provided multiple oral reading responses (Loukina et al., 2019). The performance of the engine is comparable to existing automatic ORF scoring engines such as FLORA (Bolaños et al., 2013) and Moby.Read (Bernstein et al., 2017).

Analysis

To address RQ1 on the range of readability of book passages, we examined the distribution of grade levels of book passages. To address RQ2 and RQ3, we fit a series of random effects models using the *lme4* package (Bates et al., 2015) in *R*, the statistics software. We used the maximum likelihood method for model estimation and compared model fit across models with different fixed and random effects. The data set was organized as a data table in a long format, with each row representing an individual reading a HP passage. This data table had 6,092 rows, which consisted of each of the 50 students contributing between 25 and 297 oral responses, with the median at 107 oral responses.

To address RQ2 exploring sources of variation in WCPM (i.e., person, text, vs. residual), we fit an unconditional model (Model 1) with random intercepts at person and text levels, along with random residual errors. The WCPM of a student i reading aloud passage j is denoted as:

$$\text{WCPM}_{ij} = \beta_0 + u_{0i} + u_{0j} + \epsilon_{ij} \qquad \text{(Model 1)}$$

As such, the WCPM of a student reading a passage was determined by characteristics of the student, the passage, and other unspecified factors. By treating both student and passage as random effects, we were able to estimate how variance in students' WCPM was distributed across students versus passages.

To address RQ3, we added the ongoing amount of book reading to the random effects model to examine changes in students' ORF as they read aloud more of the book (Model 2). Specifically, we derived the variable Book Progress for each passage, defined as the number of words (10,000 words as the unit) of the book before the passage. This variable was treated as a fixed effect in Model 2, assuming every student had the same growth rate.

$$\text{WCPM}_{ij} = \beta_0 + \beta_1 \cdot \text{book\_progress}_{ij} + u_{0i} + u_{0j} + \epsilon_{ij} \qquad \text{(Model 2)}$$

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To examine whether some students had excessive influence when estimating the effect of book progress, resulting in unstable estimates, we treated students as a fixed effect in Model 3 and compare the coefficients of book progress in Models 2 and 3.

$$WCPM_{ij} = \beta_0 + \beta_1 \cdot book\_progress_{ij} + \beta_2 \cdot person_i + u_{0j} + \epsilon_{ij} \quad \text{(Model 3)}$$

Finally, we added a random effect of Book Progress to the model to allow the growth rate to vary between students in Model 4.

$$\text{WCPM}_{ij} = \beta_0 + \beta_1 \cdot \text{book\_progress}_{ij} + u_{0i} + u_{1i} \cdot \text{book\_progress}_{ij} + u_{0j} + \epsilon_{ij} \quad \text{(Model 4)}$$

Results

The HP book has about 77 thousand words. The average location of the last student oral reading response from the sample was 75 thousand words, $SD = 4$ thousand words, and the range was between 61 and 77 thousand words. All students provided oral reading responses covering at least 80% of the book, and 90% students continued to provide oral reading responses after 90% of the book. Collectively, our student sample produced 6,092 scorable oral reading responses, which consisted of 1,093 unique passages. Addressing RQ1 on the comparability of book passages in terms of estimated grade level, the average length of passages read aloud by students was 109 words, $SD = 44$ words, or 9 sentences, $SD = 4$ sentences. The mean grade level as estimated by TextEvaluator (Sheehan et al., 2014) was Grade 5, $SD = 2$ grade levels, and the range was from Grade 2 to Grade 12 (Figure 1). About 75% of all passages were between Grade 3 and Grade 7, inclusive.

A series of random effects models were conducted to address RQ2 and RQ3. RQ2 concerns the sources of variation in WCPM when children read the book. For this, we looked at the variance components of the random effects of Model 1. Results showed that 56% (variance = 434) of the variance in WCPM scores was between students, 13% (variance = 100) was between passages, and 32% (variance = 246) was within students (Table 1). The fixed effect of Model 1 (β_0) shows that students' average WCPM when reading the book was 102 WCPM.

To address RQ3, Model 2 included book progress as a predictor to examine how reading more words in the book contributed to changes in WCPM. The fixed effects of Model 2 indicated that students on average started reading the book at 93 WCPM (β_0), and with each additional 10,000 words of reading their ORF increased by 2.35 WCPM (β_1). Adding book progress as a predictor significantly improved model fit compared to Model 1, $\chi^2(1) = 190$, $p < .01$ (see Table 1 for model deviance).

As a precaution, we also explored Model 3 in which student was treated as a fixed effect (rather than random effect as in Models 1 and 2). After controlling for student characteristics, the slope for book progress was 2.35 WCPM per 10,000 words read. This value was the same compared to Model 2 in which student was treated as a random effect. The convergence of results in the two models suggests a low risk of bias due to participant idiosyncrasies when estimating the effect of book progress.

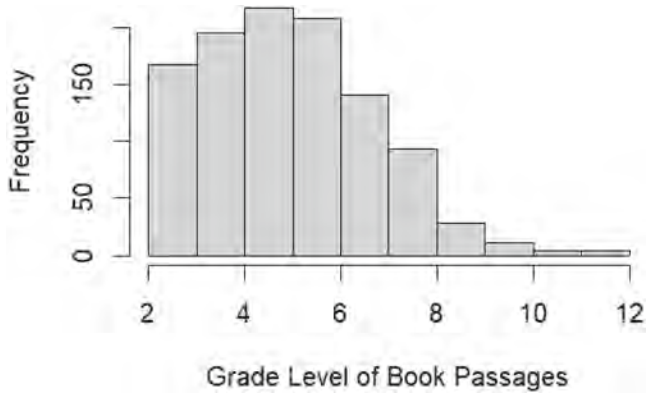

Figure 1 Distribution of Grade Level of Book Passages Read Aloud by Students

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1 Variance Components and Model Fit of Random Effects Models Predicting Word-Correct-Per-Minute

Component	Model 1	Model 2	Model 3	Model 4
Passage	100	70	70	70
Student	434	434		438
Growth				3.44
Residual	246	247	247	231
Model Deviance	52,170	51,980	51,677	51,719
Number of Parameters	4	5	53	7

*Growth is represented by per 10,000 words of book progress.

Finally, we included book progress as a random effect in Model 4 to examine the between-individual variation in WCPM growth. The fixed effects of Model 4 showed that students started reading the book at 93 WCPM, and each 10,000 words of reading was associated with improvement of 2.26 WCPM. Random effects showed that the variance of book progress was 3.44. In other words, the standard deviation of growth rate as measured by book progress was 1.85 WCPM. Adding random effects associated with book progress significantly improved model fit compared to Model 2, $\chi^2(2) = 261.6$, $p < .01$.

Discussion

The goal of this study was to explore whether we can detect changes in children's ORF as then engaged in sustained reading of a book. We adopted a widely used progress monitoring measure, ORF, and applied it to the context of electronic shared book reading using Relay Reader. Students used Relay Reader and alternated between listening (and reading along) and reading aloud themselves. We used automatic speech recognition to monitor WCPM as an indicator of ORF. Because the book passages are not standardized as passages in traditional ORF assessments, we first analyzed the distribution of readability of book passages.

RQ1 What is the range of readability of passages in the book, and how does it compare to the range of readability in standardized ORF assessments?

Traditional ORF measures constrain text readability of probe passages in an effort to make them "equivalent" (Ardoin et al., 2005; Francis et al., 2008; Good et al., 2013). In contrast, the book passages that students read aloud in this study were not controlled for readability and they displayed a wide range of variation. The TextEvaluator grade level of book passages ranged between grade 2 to 12 (Figure 1). This is larger than the range of grade levels of passages used in standardized ORF assessments. For example, Toyama et al. (2017) analyzed the grade level of passages used in DIBELS, also with TextEvaluator, and found that DIBELS passages used in the same grade level on average had a range of 3 grade levels, and the widest range was 5 grade levels (Grade 4 passages; p. 156). If we use the range of 3 grade levels as a criterion, about one third of book passages were outside of this range; if we use the range of 5 grade levels as a criterion, about 20% of book passages were outside of this range. Thus, the book passages had wider range in readability than passages used in traditional ORF assessments like DIBELS.

Prior research suggests that nonstandardized passages may add to unexplained variation in WCPM scores (Christ et al., 2012; Wayman et al., 2007), creating imprecisions when estimating growth rate. RQ2 was aimed at understanding the sources of variation in students' WCPM scores.

RQ2 How much variation in students' WCPM is contributed by text differences vs. individual differences?

According to the random effects of Model 1, 13% of the variance in WCPM scores was associated with differences in passages. This is consistent with prior research. Francis et al. (2018) collected longitudinal data of over 1,500 Grade 6–8 students reading aloud multiple, independent passages that had a wide range of text features. They found that 11% of variance in WCPM scores was from passages. Additionally, Model 1 showed that the largest contributor to variation in WCPM scores was individual differences, which is also consistent with Francis et al. (2018). These results suggest that the contribution of individual differences to WCPM likely overwhelms that of text difference to WCPM, supporting the usage of more diverse texts in accessing students' ORF for benchmarking purposes.

However, compared to Francis et al. (2018), the proportion of variance at the person level was smaller (56% vs. 73%), and the proportion of variance at the within-person level was larger (32% vs. 16%). Several factors including the student

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

samples, diversity of texts, and assessment contexts might have contributed to these differences. Importantly, the within-person variance in WCPM scores contains growth effects, which was further examined in RQ3.

RQ3 Do students' WCPM improve as they read more of the book?

Model 4 results showed that (a) children's WCPM when they started reading the book was about 93 WCPM and (b) reading each 10,000 words of the book corresponded to 2.26 WCPM improvement. How does this growth rate compare to existing literature? Because previous studies have only focused on the growth rate by time, it is necessary to convert this growth rate to time-based estimates.

Based on the growth rate, an average student who finished reading the 77,000-word book is expected to improve ORF by $7.7 \times 2.26 = 17$ WCPM. The average time students finished reading the book was about 10 weeks. Therefore, assuming linear growth, it can be estimated that students when reading the HP book on Relay Reader improved their ORF by 1.7 WCPM per week.

Deno et al. (2001) analyzed a database that included about 3,000 longitudinal ORF data for students in Grades 1–6. They estimated that fourth-grade general education students on average had weekly gain of 1.01 WCPM. Deno et al. (2001) also estimated growth rate in ORF separately by students' initial WCPM. For students who were comparable to the sample in this study in initial WCPM (i.e., between 81–110 in Deno et al., 2001), their growth rate was 1.36 WCPM.

In an intervention study comparing the effects of repeated reading vs. wide reading, Ardoin et al. (2016) randomly assigned 168 Grade 2 students to one of three conditions: repeated reading, wide reading, or business as usual. Students in the first two groups received reading interventions four times per week for about 10 weeks. Students' WCPM was monitored with a weekly ORF test using passages independent of the materials used in the intervention. Results showed that depending on reading level, students in the repeated reading condition gained between 1.19 and 2.10 WCPM per week, students in the wide reading condition gained between 1.47 and 2.10 WCPM per week, and students in the business-as-usual condition gained between 0.77 and 1.47 WCPM per week. The growth rate we estimated from the current study (1.7 WCPM/week) was within the range of these prior studies (Ardoin et al., 2016; Deno et al., 2001).

While the growth rate of the current study is in line with prior research, there are other notable findings. For instance, an analysis of random effects of Model 4 revealed substantial variation between individuals in growth. The standard deviation of the growth rate estimate between students was 1.85 WCPM. Assuming normal distribution, this means that about 16% students' growth rate was lower than .4 WCPM per 10,000 words and another 16% was above 4 WCPM per 10,000 words. This significant individual difference in growth rate in WCPM suggests not all students' ORF improved equally with the book reading activity. Identifying students who are not making adequate progress and providing additional support may help improve the effectiveness of book reading on ORF growth.

Implications for Reading Assessment

An interesting finding of this study is that, despite book passages having a much wider range of readability compared to passages used in standardized ORF assessments, differences between passages only explained a small portion of variance in students' WCPM scores (13%). This is consistent with Francis et al. (2018), who found text level differences explained 11% of variance in students' WCPM scores. The implication is that the evaluation of students' ORF perhaps does not have to be limited to passages with a narrow range of text features (Barth et al., 2014; Francis et al., 2008; Toyama et al., 2017). Passages with a wider range of text features *can* still distinguish students' ORF.

One consideration in monitoring ORF from children's oral reading of book passages is the quality of data for detecting growth. Christ et al. (2012) established standards on the quality of ORF progress monitoring data based on a review of empirical studies. Specifically, the residual variance of WCPM, denoted as σ_ϵ^2, ranged between 25 and 400 for data sets with quality ranging from "very good" to "very poor." In comparison, the variance component of Model 4 showed that the residual variance was 231, which would be classified as "poor" according to Christ et al.'s (2012) standard. This is not surprising given the low stakes, nonstandardized, and unobtrusive assessment context: We merely observed students' oral reading as they read a full-length book instead of giving them a test. Nevertheless, the data set revealed significant growth in students' WCPM as they read the book on Relay Reader, thanks to the larger number of individual observations (median = 107 oral reading responses) compared to traditional progress monitoring studies. Increasing the number of observations helps reduce the level of standardization required of the assessment development process, thus allowing more flexible use of ORF as a progress monitoring measure.

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Limitations and Future Directions

First, future research should investigate how much the ORF gain we observed when children read the book transfers to other contexts. One way to address this question is to provide more books for children to read on Relay Reader. Potential transfer of ORF improvement could be evaluated when children finish reading one book and move on to another. Additionally, future research should explore whether other modes of reading (e.g., silently, varying amount of listening, etc.) also lead to similar gains in fluency.

Second, although we have collected many oral responses from each student, the number of students in this study was relatively small compared to the number of book passages. On average, each book passage was only read aloud by about 5-6 students. The relatively small number of oral responses collected for each unique passage in this study might have limited our precision when estimating passage effects on WCPM. The precision of our results will likely be improved with each passage read aloud by more students.

Third, it remains to be explored what explains the ORF improvement that we observed in this study. One such division is accuracy versus speed. How much of the WCPM increase can be attributed to better word reading accuracy (i.e., to learn unfamiliar words and read them more accurately in later parts of the book) as opposed to improved articulation speed and prosody? Another division is the content knowledge of the book versus the automaticity of reading. How much of the WCPM increase is due to students' better knowledge of HP and improved mental model of the plot as they read more? How much of the WCPM increase is due to more automatized reading processing? Again, we believe this question can be answered with an examination of students' ORF changes across books. A larger ORF transfer as the child starts reading a new book would be evidence supporting the automaticity component. However, we believe the accumulation of content knowledge may not be entirely separable from the improvement of one's reading skills, since the development of reading skills depends on one practicing reading by reading of materials on some content areas.

Finally, in addition to book reading, other factors such as instruction, natural maturation, or other language and reading activities may also contribute to the observed improvement in students' ORF. We leave it to future research to examine the unique contribution of book reading to ORF growth through experimental studies.

Conclusion

This study demonstrates how children's WCPM change as they read a book on Relay Reader, an electronic shared book reading platform. Results underscore the potential of ORF as a progress monitoring measure in the context of electronic shared book reading. Despite the wide range of readability in the book passages—far greater than passages used in standardized ORF assessments—a sufficient number of oral reading responses afforded by this context enabled us to detect significant improvement in WCPM scores as students read more of the book.

Acknowledgments

We'd like to thank Jonathan Schmidgall, Lin Gu, Ikkyu Choi, Dan McCaffrey, and many others for their feedback on earlier versions of this manuscript.

References

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, *233*(3), 285–303. https://doi.org/10.1598/RRQ.23.3.2

Ardoin, S. P., Binder, K. S., Foster, T. E., & Zawoyski, A. M. (2016). Repeated versus wide reading: A randomized control design study examining the impact of fluency interventions on underlying reading behavior. *Journal of School Psychology*, *59*, 13–38. https://doi.org/10.1016/j.jsp.2016.09.002

Ardoin, S. P., Suldo, S. M., Witt, J., Aldrich, S., & McDonald, E. (2005). Accuracy of readability estimates predictions' of CBM performance. *School Psychology Quarterly*, *20*(1), 1–22. https://doi.org/10.1521/scpq.20.1.1.64193

Baker, L., Fernandez-Fein, S., Scher, D., & Williams, H. (1998). Home experiences related to the development of word recognition. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 263–287). Lawrence Erlbaum Associates.

Baker, L., Scher, D., & Mackler, K. (1997). Home and family influences on motivations for reading. *Educational Psychologist*, *32*(2), 69–82. https://doi.org/10.1207/s15326985ep3202_2

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Barth, A. E., Tolar, T. D., Fletcher, J. M., & Francis, D. (2014). The effects of student and text characteristics on the oral reading fluency of middle-grade students. *Journal of Educational Psychology*, *106*(1), 162–180. https://doi.org/10.1037/a0033826

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Beigman Klebanov, B., Loukina, A., Madnani, N., Sabatini, J., & Lentini, J. (2019). Would you? Could you? On a tablet? Analytics of children's ebook reading. In S. Hsiao, J. Cunningham, K. McCarthy, G. Lynch (Eds.), *LAK19: Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 106–110). ACM. https://doi.org/10.1145/3303772.3303833

Beigman Klebanov, B., Loukina, A., Sabatini, J., & O'Reilly, T. (2017). Continuous fluency tracking and the challenges of varying text complexity. In J. Tetreault, J. Burstein, C. Leacock, & H. Yannakoudakis (Eds.), *Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 22–32). ACL. https://doi.org/10.18653/v1/W17-5003

Bernstein, J., Cheng, J., Balogh, J., & Rosenfeld, E. (2017, August 20-24). Studies of a self-administered oral reading assessment. *7th ISCA Workshop on Speech and Language Technology in Education*, Stockholm, Sweden. https://www.isca-archive.org/slate_2017/bernstein17_slate.pdf

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*(2), 294–304. https://doi.org/10.1037/a0013844

Bolaños, D., Cole, R. A., Ward, W. H., Tindal, G. A., Hasbrouck, J., & Schwanenflugel, P. J. (2013). Human and automated assessment of oral reading fluency. *Journal of Educational Psychology*, *105*(4), 1142–1151. https://doi.org/10.1037/a0031479

Bus, A. G., Van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*(1), 1–21. https://doi.org/10.3102/00346543065001001

Carver, R. P., & Leibert, R. E. (1995). The effect of reading library books at different levels of difficulty upon gain in reading ability. *Reading Research Quarterly*, *30*(1), 26–48. https://doi.org/10.2307/747743

Christ, T. J., Zopluoglu, C., Long, J. D., & Monaghen, B. D. (2012). Curriculum-based measurement of oral reading: Quality of progress monitoring outcomes. *Exceptional Children*, *78*(3), 356–373.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*(3), 219–232. https://doi.org/10.1177/001440298505200303

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, *30*(4), 507–524. https://doi.org/10.1080/02796015.2001.12086131

Eberwein, L. D. (1979). The variability of readability of basal reader textbooks and how much teachers know about it. *Reading World*, *18*(3), 259–272. https://doi.org/10.1080/19388077909557484

Esteves, K. J., & Whitten, E. (2011). Assisted reading with digital audiobooks for students with reading disabilities. *Reading Horizons*, *51*(1), 21–40.

Francis, D. J., Kulesz, P. A., & Benoit, J. S. (2018). Extending the simple view of reading to account for variation within readers and across texts: The complete view of reading (CVR i). *Remedial and Special Education*, *39*(5), 274–288. https://doi.org/10.1177%2F0741932518772904

Francis, D. J., Santi, K. L., Barr, C., Fletcher, J. M., Varisco, A., & Foorman, B. R. (2008). Form effects on the estimation of students' oral reading fluency using DIBELS. *Journal of School Psychology*, *46*(3), 315–342. https://doi.org/10.1016/j.jsp.2007.06.003

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256. https://doi.org/10.1207/s1532799xssr0503_3

Good, R., Kaminski, R., Dewey, E., Wallin, J., Powell-Smith, K., & Latimer, R. (2013). *DIBELS Next technical manual*. Dynamic Measurement Group.

Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, *15*(1), 75–90. https://doi.org/10.1016/S0885-2006(99)00038-1

Harlaar, N., Deater-Deckard, K., Thompson, L. A., DeThorne, L. S., & Petrill, S. A. (2011). Associations between reading achievement and independent reading in early elementary school: A genetically informative cross-lagged study. *Child Development*, *82*(6), 2123–2137. https://doi.org/10.1111/j.1467-8624.2011.01658.x

Hasbrouck, J., & Tindal, G. (2017). *An update to compiled ORF norms* (Technical Report No. 1702). Behavioral Research and Teaching, University of Oregon.

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, *59*(7), 636–644. https://doi.org/10.1598/RT.59.7.3

Ingham, J. (1982). *Books and reading development: The Bradford book flood experiment*. Heinemann Educational Books.

Kim, J. S., & Guryan, J. (2010). The efficacy of a voluntary summer book reading intervention for low-income Latino children from language minority families. *Journal of Educational Psychology*, *102*(1), 20–31. https://doi.org/10.1037/a0017270

Liceralde, V. R. T., Loukina, A., Beigman Klebanov, B., & Lockwood, J. R. (2022). Beyond text complexity: Production-related sources of text-based variability in oral reading fluency. *Journal of Educational Psychology*, *114*, 16–36. https://doi.org/10.1037/edu0000532

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Loukina, A., Beigman Klebanov, B., Lange, P., Qian, Y., Gyawali, B., Madnani, N., Misra, A., Zechner, K., Wang, Z., & Sabatini, J. (2019). Automated estimation of oral reading fluency during summer camp e-book reading with MyTurnToRead. *9th ISCA Workshop on Speech and Language Technology in Education*, Stockholm, Sweden.

McBride-Chang, C., Manis, F. R., Seidenberg, M. S., & Custodio, R. G. (1993). Print exposure as a predictor of word reading and reading comprehension in disabled and nondisabled readers. *Journal of Educational Psychology*, *85*(2), 230–238. https://doi.org/10.1037/0022-0663.85.2.230

Mol, S. E., & Bus, A. G. (2011). To read or not to read: a meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267–296. https://doi.org/10.1037/a0021890

Powell-Smith, K. A., Good III, R. H., & Atkins, T. (2010). *DIBELS Next oral reading fluency readability study (Technical Report No. 7)*. Dynamic Measurement Group.

Renaissance Learning. (2018). *Education Leader's Guide to Reading Growth*. Renaissance Learning, Inc.

Renaissance Learning. (2019). *What kids are reading: World's largest annual study of K–12 reading habits: 2019 edition*. https://www.renaissance.com/wkar-report/

Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, *46*(3), 343–366. https://doi.org/10.1016/j.jsp.2007.06.006

Rowling, J. (1999). *Harry Potter and the sorcerer's stone*. Arthur A. Levine Books.

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, *14*(3), 245–302. https://doi.org/10.1006/drev.1994.1010

Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *The Elementary School Journal*, *115*(2), 184–209. https://doi.org/10.1086/678294

Topping, K. (1987). Paired reading: A powerful technique for parent use. *The Reading Teacher*, *40*(7), 608–614.

Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M. K., Tolvanen, A., & Poikkeus, A. M. (2020). Leisure reading (but not any kind) and reading comprehension support each other—A longitudinal study across Grades 1 and 9. *Child Development*, *91*(3), 876–900. https://doi.org/10.1111/cdev.13241

Toyama, Y., Hiebert, E. H., & Pearson, P. D. (2017). An analysis of the text complexity of leveled passages in four popular classroom reading assessments. *Educational Assessment*, *22*(3), 139–170. https://doi.org/10.1080/10627197.2017.1344091

Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, *41*(2), 85–120. https://doi.org/10.1177%2F00224669070410020401

What Works Clearinghouse. (2015). *WWC intervention report: Shared book reading*. U.S. Department of Education & Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/EvidenceSnapshot/458

Suggested citation:

Wang, Z., Beigman Klebanov, B., O'Reilly, T., & Sabatini, J. (2024). *Monitoring oral reading fluency from electronic shared book reading: Insights from a full-length book reading study with Relay Reader®* (Research Report No. RR-24-14). ETS. https://doi.org/10.1002/ets2.12390

Action Editor: Jonathan Schmidgall and Daniel F. McCaffrey

Reviewers: Ikkyu Choi and Lin Gu

ETS, the ETS logo, RELAY READER, and TEXTEVALUATOR are registered trademarks of Educational Testing Service (ETS). All other trademarks are property of their respective owners.

Find other ETS-published reports by searching the ETS ReSEARCHER database.

23308516, 2024, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ets2.12390, Wiley Online Library on [30/01/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License