



Research Memorandum

ETS RM-24-01

Exploring the Use of the Relay Reader™ App With Young English as Foreign Language Learners: A Case Study

Veronika Timpe-Laughlin
Beata Beigman Klebanov
Anastassia Loukina

January 2024



ETS Research Memorandum 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

Heather Buzick
Senior Research Scientist

Tim Davey
Director Research

Larry Davis
Director Research

Jamie Mikeska
Senior Research Scientist

Jonathan Schmidgall
Senior Research Scientist

Jesse Sparks
Senior Research Scientist

Klaus Zechner
Senior Research Scientist

PRODUCTION EDITORS

Kim Fryer
Manager, Editing Services

Ayleen Gontz
Senior Editor

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 Memorandum 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 Memorandum 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.

**Exploring the Use of the Relay Reader™ App With Young
English as Foreign Language Learners:
A Case Study**

Veronika Timpe-Laughlin, Beata Beigman Klebanov, and Anastassia Loukina
Educational Testing Service, Princeton, New Jersey, United States

January 2024

Corresponding author: V. Timpe-Laughlin, E-mail: vlaughlin@ets.org

Suggested citation: Timpe-Laughlin, V., Beigman Klebanov, B., & Loukina, A. (2024). *Exploring the use of the Relay Reader™ app with young English as foreign language learners: A case study* (Research Memorandum No. RM-24-01). ETS.

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

To obtain a copy of an ETS research report, please visit
<https://www.ets.org/contact/additional/research.html>

Action Editor: Jonathan Schmidgall

Reviewers: Ching-Ni Hsieh and Shinhye Lee

Copyright © 2024 by Educational Testing Service. All rights reserved. ETS, the ETS logo,
TEXTEVALUATOR, TOEFL, and TOEFL JUNIOR are registered trademarks of Educational Testing Service (ETS).

RELAY READER is a trademark of ETS. All other trademarks are the property of their respective owners.

Abstract

Given the pandemic-induced mass transition to online learning worldwide, instructional practices are being adjusted to new realities such as reduced schedules and increased reliance on technology. It is therefore important to develop an understanding of instructional practices that are well suited for such contexts. In this study, we aimed to evaluate the feasibility of using Relay Reader™, a digital application for reading books and other printed texts, to implement a technology-based supplemental extensive reading program in a fully remote English as a foreign language high school context using data collected at a high school in Chile. The detailed case study suggests the approach is highly feasible and highlights the importance of understanding the digital learning preferences of different students.

Keywords: COVID-19, Relay Reader™, TOEFL Junior® Standard, digital app, reading, learning preference, technology in distance learning, English as a foreign language (EFL), Chile, secondary education

Acknowledgments

Anastassia Loukina is currently at Grammarly.

Table of Contents

Method	4
Study Context.....	6
Participants	7
Instruments.....	8
TOEFL Junior Test.....	8
Data Collection Procedure	9
Data Analysis.....	11
Results.....	12
Discussion.....	25
Research Question 1: To what extent do students engage with the Relay Reader application?.....	25
Research Question 2: What are the teacher’s and students’ perceptions of the technology-mediated ER practice, in particular, with regard to feasibility of implementation, complexity of the reading passage, and perceived student motivation?.....	26
Concluding Remarks.....	28
References	30
Appendix A. Relay Reader Teacher Protocols.....	34
Appendix B. Data Preparation for the Analysis of Longitudinal Trends	38
Notes.....	40

Reading in the target language is one of the key skills that second and foreign language learners need to acquire and practice. Students typically learn to read by engaging in reading activities. A particular form of reading practice is extensive reading (ER) or uninterrupted sustained silent reading. Grabe (2009) noted that ER “requires each student to read a book at a designated time for a period (usually 15–20 minutes) that is not interrupted for any purpose” (p. 327). Additionally, Nation (2008) highlighted that during ER, learners “should be reading with their attention on the meaning of the text rather than on learning the language features of the text” (p. 50). In other words, ER involves reading longer texts for enjoyment either inside or outside of the second/foreign language (L2) classroom with the primary goal of developing general L2 reading skills.

Research has shown that ER can have several benefits for first language (L1) and L2 learners, including gains in vocabulary, reading comprehension, and other literacy skills (Elley, 1992, 2001; Elley & Mangubhai, 1983; Grabe, 2009; Guthrie et al., 2001; Horst, 2005; Nakanishi, 2015; Nation, 2001; Schmitt, 2008). Although ER can only occur if about 95% of the words in a text are already familiar to the learner or at least do not constitute a burden to the learner (Hsueh-chao & Nation, 2000), studies in the context of L2 instruction have found that vocabulary tends to be learned every time a learner encounters a new word while reading (Nation, 2008). With regard to reading comprehension, research shows that amount of reading is associated with reading comprehension achievements (Grabe, 2009) and, in particular, reading fluency (Perfetti, 2007). Additionally, Tsang (1996) and Hafiz and Tudor (1989) demonstrated that ER in the target language improves L2 writing skills. Hence, several studies have reported positive results associated with ER, including proficiency gains in various language-related skills.

In addition to language-related benefits, ER has been found to be more motivating than traditional textbook-oriented reading instruction (Cox & Guthrie, 2001; Day & Bamford, 1998; Grabe, 2009; Guthrie et al., 2001). For example, Day and Bamford (1998) reviewed ways to promote ER in the classroom. A particular aspect they highlighted as motivating for students was a focus on reading-fluency activities, including having the teacher read aloud to the students and multiple re-reading of the text. The motivational element of ER and the promotion

of fluency may be rooted in how the teacher sets up the reading experience as well as in the text types. For instance, research has pointed out that the dyadic read aloud in collaboration with an experienced reader provides a scaffolding structure and guidance that may boost motivation among less proficient readers (Mehigan, 2020). Additionally, in the context of ER experiences, students tend to read authentic texts rather than short, often decontextualized reading passages tailored to L2 learners that are commonly found in textbooks (Siegel, 2014). These aspects may increase motivation for reading and, by extension, fluency, which tends to be an artefact of the amount of reading that students engage in.

However, despite the widely documented evidence in favor of ER, there are relatively few carefully controlled experimental training studies, in particular in L2 reading contexts (Grabe, 2009). Whereas many ER studies have focused on L1 readers, only a few studies have investigated L2 learners, with even less research carried out with *young* L2 readers¹ (Tanaka & Stapleton, 2007). For example, Nakanishi's (2015) meta-analysis showed a "lack of younger participants in ER research" (p. 30). He noted that more than half (56%) of the participants in studies that investigated ER in the L2 were university-level students. Although his findings show greater effects of ER in older learners, he argued that "it could also be more valuable for younger participants; studies (e.g., Hafiz & Tudor, 1989; Tsang, 1996) have shown that it might improve reading proficiency and develop the joy of reading at an earlier stage of learning" (Nakanishi, 2015, p. 31)—a line of research we would like to expand with this study.

Among the few studies carried out with young learners, several studies have focused on out-of-school ER reading, using reading logs or questionnaires to investigate students' ER experiences and activities. For example, in the United States, Anderson et al. (1988) examined the reading logs about daily out-of-school reading activities that 155 fifth graders kept for durations ranging from 8 to 26 weeks. They controlled for students' reading abilities in Grade 2 and found that the amount of extracurricular reading, especially the reading of books, was the best predictor for gains in reading achievements. Similarly, Tanaka and Stapleton (2007) conducted a study of English as a foreign language (EFL) ER with 96 Japanese high school students for a period of 5 months. They included more of an in-class focus insofar as teachers read short reading passages to their students during the first 5–10 minutes of each class.

Additionally, they clarified unknown words and vocabulary. Then students were tasked to read the same passages again at home. Compared to a control group who did not receive teacher instructions or guidance while reading, the experimental group also showed higher reading-rate increases as well as better reading comprehension. Although these studies provide initial insights into the implementation of ER in the context of the EFL classroom in general, we are not aware of studies that focus on ER in distance or remote² EFL teaching and learning contexts, that is, in contexts that have become increasingly important since the COVID-19 pandemic where students are not learning EFL face-to-face with a teacher inside the traditional brick-and-mortar classroom.

Finally, previous studies that have investigated ER in L2 education used paper-based books, not technology-mediated texts. Research that has focused on technology in the context of L2 reading has mainly examined the use of support features such as concordances, hypertext glosses, and e-dictionaries to aid vocabulary learning (Cobb *et al.*, 2001). To our knowledge, no studies exist that investigated ER experiences in which students used tablets or smartphones, tools that are used by the majority of adolescents and whose use has increased during the COVID-19 pandemic (Vogels *et al.*, 2022) to deliver reading passages.

To summarize, despite strong empirical evidence of the benefits of ER practice, Grabe (2009) noted that “it is extraordinary that extensive reading is still treated as a fringe issue in L2 reading instruction” (p. 328; see also Renandya & Jacobs, 2002). Also, given limited research in L2 ER practices as well as the increasing use of online learning tools to meet the challenges of pandemic-induced learning environments, a focus on technology-mediated ER (using tablets and smartphones—devices that are becoming increasingly prominent in L2 learners’ real-life contexts) may not only advance the knowledge base and add to previous research, but also provide a solution that could aid with the implementation of ER in L2 instruction, especially in times of remote teaching and learning.

In this study, we implemented an ER program in a fully remote EFL education context in Chile using Relay Reader™, an application that allows students to read a book. In doing so, they can take turns reading and following the text with a highly proficient English-speaking human narrator. To explore students’ reading in the application as well as their and their teacher’s

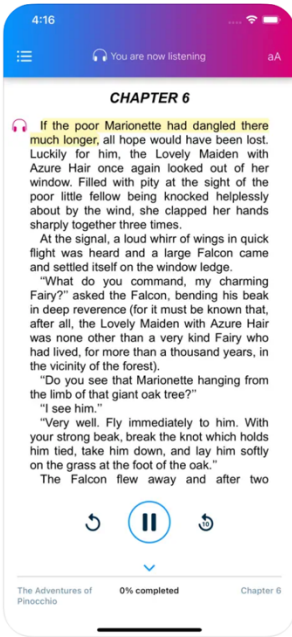
opinions about the technology-mediated reading experience, we put forth the following two research questions:

1. To what extent do students engage with the Relay Reader application?
2. What are the teacher's and students' perceptions of the technology-mediated ER practice, in particular with regard to the feasibility of implementation, complexity of the reading passage, and perceived student motivation?

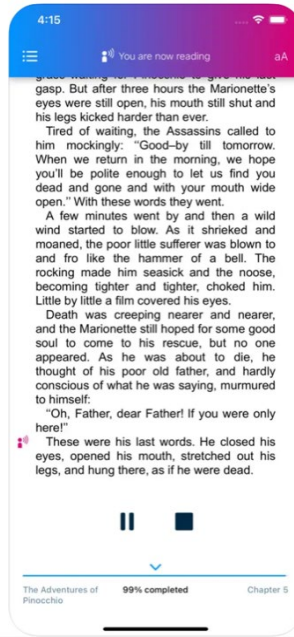
Method

To explore the research questions, we used the Relay Reader app (<https://relayreader.org/>; Madnani et al., 2019), a reading and listening application that allows readers to read a book, taking turns with a highly proficient English-speaking human narrator. The app is available as a web-based or mobile app (see Figure 1). It featured the book *The Adventures of Pinocchio* by Carlo Collodi translated by Della Chiesa, a 130-page 39,000-word unabridged literary classic.³ When readers interact with the app, they can listen to the human narrator, following on the screen as text is highlighted when it is read (see Figure 1, a). Then they read aloud their passage (see Figure 1, b). The default setting is an alternation between 200 words read by the narrator and 150 words read by a user (see Figure 1, c). The app has several features that allow readers to (a) replay and listen to the narrator read the same passage or parts of it several times, (b) listen to their own reading, (c) listen to the narrator read their passage, and (d) re-record themselves reading their passage multiple times (see Figure 1, d). Additionally, readers are frequently asked comprehension questions about the story that are intended to help them stay focused (see Figure 1, f)—readers respond to two questions after every other reader turn. For the default setting described above, this corresponds to about two questions per 700 words of running text or, roughly, one question per page. Finally, readers can access their reading history that shows progress in terms of how much they read, how far they have progressed in the story, and how long they have been reading (see Figure 1, e). The app logs activity (timestamps for listening and reading turns, responses to questions), records the oral readings, and automatically generates performance indices such as accuracy (proportion of words on the page read correctly) and fluency (the number of words read correctly per minute).

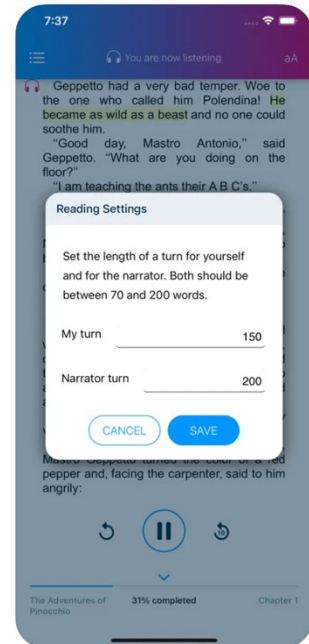
Figure 1. Relay Reader’s Design Features



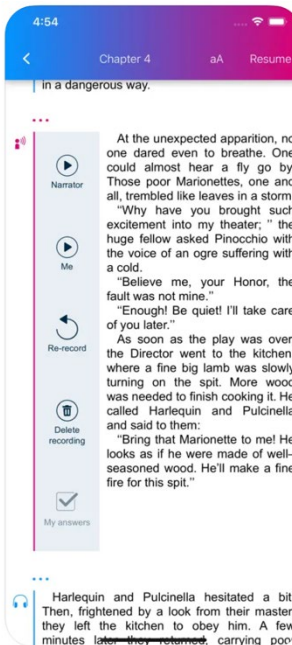
(a) Text read by narrator is highlighted



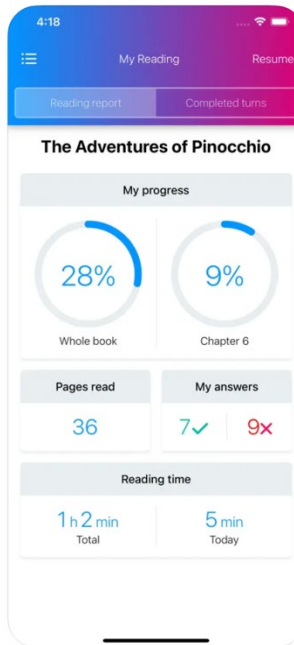
(b) Visual indicator for readers marking their respective reading passage



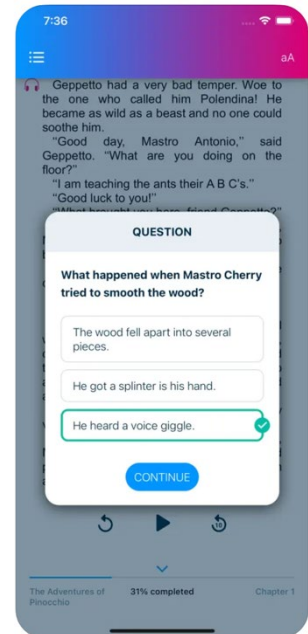
(c) Settings indicating the default length of reading turns



(d) Replay and re-record functions



(e) Reading report



(f) Comprehension question

In this case study, fluency estimates and comprehension (percentage of comprehension items answered correctly) were shared with the teacher in weekly reports. Both accuracy and the fluency measures are generated by the automated engine developed in house (Loukina *et al.*, 2019).

The Relay Reader app was selected for this study because the app (a) can be used with an ER approach and (b) constitutes a freely available tool that was designed by ETS research and is therefore in need of empirical backing. The design of the Relay Reader app allows for the app to be used for an ER approach insofar as it includes a whole book and several features that facilitate aspects inherent to ER. For example, as a book-length text, *The Adventures of Pinocchio* provides a relatively large amount of reading material that readers can engage in for joy, reading extensively rather than intensively. Additionally, the focus on comprehension of the contents (i.e., understanding the plot) is prioritized over analyzing every detail (Grabe, 2009; Nakanishi, 2015). Learners, for instance, engage in quiz-like content questions that accompany the story and consolidate understanding of the text (see Figure 1f)—thus again emphasizing a focus on comprehension of meaning rather than linguistic form.

While Relay Reader has been examined before in in-person school and camp contexts with upper elementary students in the United States (Beigman Klebanov *et al.*, 2019, 2020), it has not been investigated with students who learn EFL. Additionally, no studies have been carried out in which Relay Reader was used with an ER approach in a remote learning context. Hence, thus far, ETS has not been able to make claims about the use of Relay Reader in EFL and/or remote learning contexts. The current study represents a first step into exploring a new use case—both in terms of the population (adolescent EFL learners) and learning context.

Study Context

The study was conducted in the summer of 2020. Given the timing of the study (i.e., in the midst of global lockdowns), finding schools and teachers who were willing to participate in this study was particularly challenging given the increased workload for many teachers to shift their instruction online. As such, the call for participants for this study distributed by the TOEFL® program in two countries in Latin America was answered by only two teachers. Of the

two teachers, only one was able to secure permission from school leadership to participate with his students.

This teacher taught English at a high school in Chile. Due to the COVID-19 pandemic, instruction was provided remotely. Following national and local policies and out of concern for the students' well-being, the number of English lessons per week was reduced considerably, resulting in a 1-hour-long weekly English lesson for each class. The teacher's instructional goal with the ER program was to provide supplemental reading practice to selected students whose English proficiency was deemed high enough to be able to read a literary book in English.

To that end, the teacher had invited 23 students with a relatively high English proficiency and general interest in reading from across four grades (9–12) to participate in the ER program. Students were incentivized by means of extra credit if they finished the book by the end of the study. There were no performance requirements to obtain the extra credit other than finishing the book. The teacher recommended about 1 hour of reading per week, but students were not held to any schedule. The reading activity was implemented as an out-of-class supplement; no instructional time was dedicated to discussion or review, although communication of any issues or questions regarding the reading practice was encouraged through the class chat. The teacher also made it clear that their reading activity in the Relay Reader app was monitored; that is, he would be provided with reports about their reading progress and interaction with the app.

Participants

Overall, 18 students (10 females, 8 males) and their teacher from a small private school in Chile participated in the study. The teacher was male, between 25 and 34 years old, and had 7 years of experience teaching English to secondary level learners. Of the 23 students invited to participate, five failed to secure parental consent and were excluded from the study. The students' average age was 16.1 years, ranging from 13 to 18 years of age. All participants reported Spanish as their first and home language, one additionally listed English. Two participants reported having lived in an English-speaking country. Students varied in their English learning experience—some started as early as preschool, others as late as fifth grade.

Regarding reading experience, five participants reported having read *The Adventures of Pinocchio* in Spanish. Only one student had previously listened to an audiobook in English.

No monetary compensation was provided to the site, the teacher, or the students; the reading opportunity with the new Relay Reader technology for the students and the ability to monitor the reading process by the teacher, both provided free of charge, were the benefits for the participants.

Instruments

In addition to engaging with the Relay Reader app, data were collected using the following instruments: the TOEFL Junior® Standard test, semi-structured teacher interviews, and teacher and student surveys.

TOEFL Junior Test

The TOEFL Junior Standard test measures the English language proficiency of students in middle and high school. It consists of three parts, a listening comprehension section that assesses the ability to listen and understand academic and nonacademic oral texts, a language form and meaning section that measures grammar and vocabulary knowledge, and a reading comprehension section that measures learners' ability to read and comprehend texts written in English (for more details see ETS, 2020). Students took a research form of the TOEFL Junior Standard test⁴ before they began reading in the Relay Reader app to gauge their English proficiency, specifically reading comprehension. Of the 18 participants, 14 students took the TOEFL Junior Standard test. The average total score across participants was 833 ($SD = 32.51$), which corresponds to an English proficiency level of B1 (low intermediate) on the Common European Framework of Reference (CEFR) with an average reading score of 279 ($SD = 12.28$) and an average listening score of 286 ($SD = 9.38$).

Interviews

Two interviews were conducted with the teacher to gauge his experiences with and perceptions of the Relay Reader app. The interviews were conducted in English by the three researchers via the videotelephony software program Zoom. The first hour-long interview was conducted 2 weeks into the study (i.e., once most students started the reading activity),

whereas the second interview was a poststudy interview held after students had completed their ER program (for the complete interview protocols see Appendix A). Both interviews were audio-recorded.

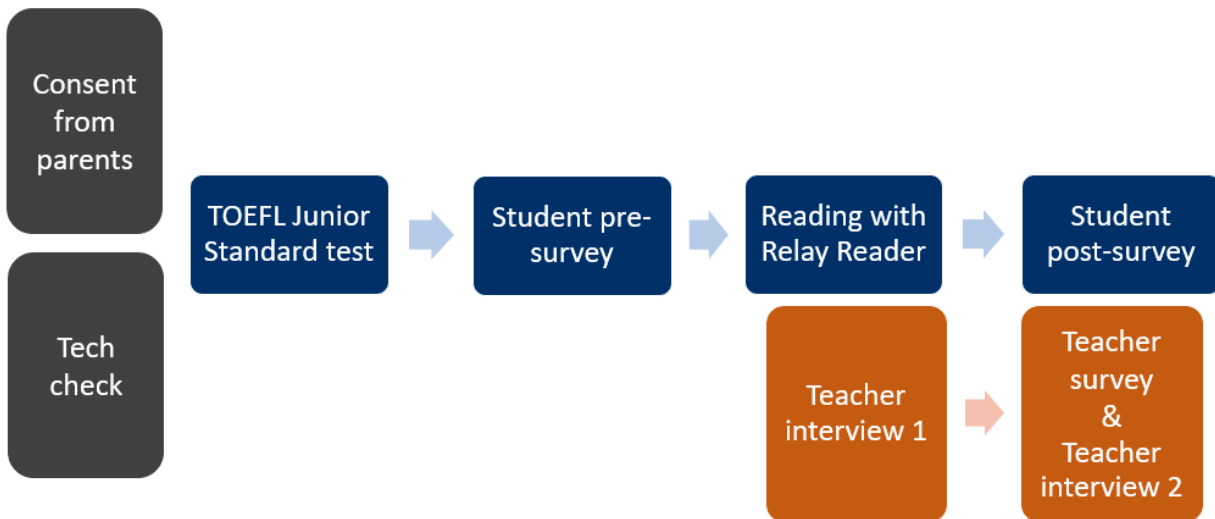
Surveys

We deployed two online surveys to students (before and after the ER practice) and one online survey to the teacher (during the ER activity). All surveys were administered in English with questions carefully worded to ensure that they would be comprehensible for students with an approximate CEFR level of B1 for reading comprehension in English. Additionally, the teacher reviewed the surveys before administration in order to provide feedback as to whether students would be able to understand the survey in English. Before the ER practice, students were asked to complete a survey that contained questions regarding the following: reading activities in L1 and L2 inside and outside of the classroom, English language learning practices, motivation regarding English learning, and demographics. At the end of the ER activity, students were asked to fill out another survey that included questions about their overall perceptions of the ER experience, their views of the app in general as well as specific technology-related aspects of it, and suggestions for improvements regarding the Relay Reader application. Additionally, students had access to optional reader experience surveys embedded in the app and available to all users—one survey after the first chapter of the book and one after the last chapter.

The teacher survey, also administered in English, included questions about perceptions of the Relay Reader app and the text, questions about supplement and support activities they used while working with the app, suggestions for improvements of the app and teaching supports, and demographic questions.

Data Collection Procedure

All data were collected between August and September of 2020. As shown in Figure 2, we began the data collection by securing parental consent and a technology check to ensure that the Relay Reader application would run smoothly on the students' devices.

Figure 2. Data Collection Process

All students then took the TOEFL Junior Standard test, and five students completed the prereading survey. Afterward, students read in Relay Reader for 6 weeks. During that time, the teacher received weekly reports from the research team about the students' reading activities. The weekly reports included the following performance indices about each student's reading activities: total time interacting with the app (including frequency counts and duration of productive and unproductive time for each student for each interaction with the app), number and pace of reading turns, number and percentage of correct responses to comprehension questions, and overall reading progress across chapters. In addition to the class-level information, the teacher was also provided with a summary of all information pertaining to a specific student. An example of a student-specific report is shown in the Figure 3.

After completing the full reading intervention, all students were asked to fill out the postreading survey. Five learners completed the survey. As students were engaged in reading, the research team conducted the first interview (60 minutes) with the teacher about 2 weeks into the study. At the end of the study, the teacher filled out a survey and a few days later participated in a final, 60-minute interview with the research team (see the interview protocols in Appendix A).

Figure 3. Student-Specific Report Example

██████P

██████P interacted with the app for 4 hours and 1 minute.

Most of ██████P's time with the app was productive: ██████P's active reading time was 1 hour and 1 minute (25%). ██████P spent 1 hour and 46 minutes (44%) on other productive activities (listening, answering questions, reviewing reading history). ██████P spent 1 hour and 14 minutes (31%) unproductively.

██████P's current reading settings are 150 words for student turn and 200 words for narrator turn.

In total, ██████P completed 63 reading turns. For all turns ██████P took the expected amount of time to read the passage.

██████P answered 55 out of 58 reading comprehension questions correctly (95%).

██████P read 76 pages from "The Adventures of Pinocchio".

Data Analysis

With regard to the extent of the students' reading activity, we first calculated the percentage of the eligible students who actually read in the application. Then, we tallied students' duration of use, the number of pages read, the number of reading turns⁵ (including the percentage of turns with clear oral reading⁶), the number of comprehension questions (including the percentage of correct answers), and the number of students who finished the book.

Additionally, we explored whether we were able to observe any patterns in students' oral reading fluency. Oral reading fluency was measured as words read correctly per minute of reading (wcpm) using the automated system built for this application (Loukina *et al.*, 2019; Misra *et al.*, 2021). In our analysis, we included the wcpm score for a student for a passage only if the student's accuracy for the passages was at least 70%. That is, we used those passages where the system had identified at least 70% of the text of the passage in the oral reading response. This cutoff is based on prior work showing that responses with low automatically measured accuracy tend to yield less reliable automated fluency scores (Beigman Klebanov & Loukina, 2021).

To explore enjoyment and perceived utility, we analyzed students' survey responses using descriptive statistics. Frequency distributions were produced and the mean and standard deviation were calculated for each survey item to determine consensus or discrepancies in opinions among students. Representative responses were extracted from open-ended survey

responses and interviews as a way of capturing and documenting response patterns in the words of the participants.

Results

In terms of the first research question—to what extent students engaged with the Relay Reader application—we found that of the 18 eligible students, 13 (72%) opened the Relay Reader application at least once over the course of the entire intervention and read with the application on average 5 hours 4 minutes, ranging from 45 minutes to 9 hours and 18 minutes. On average, students read 99 pages of the book (min = 17; max = 130) with eight students finishing the entire book (130 pages). Students completed on average 89 reading turns (min = 17; max = 143) and responded to an average of 83 reading comprehension questions (min = 14; max = 140), answering, on average, 86% of them correctly. Based on the event logs, eight of the students replayed narrator turns during reading at least once, averaging 10.6 replays (min = 1, max = 24); four of the students replayed their own recordings, averaging 10 replays (min = 1, max = 30); and four of the students re-recorded at least one of their turns, averaging 3.5 re-recordings (min = 1; max = 11). The student who replayed her own recordings 30 times also listened extensively to the narrator’s readings of her turns (14 different passages, each listened to between 1 and 4 times) and re-recorded her own reading of many of these passages (10 different passages, with one passage re-recorded twice).

Moreover, we identified a total of 921 scored oral reading turns across 13 students (average = 71, *SD* = 43, min = 0, max = 123); one reader did not produce any scorable oral readings. Table 1 shows the number of scored turns per chapter per student. Twelve of the 13 students started out reading aloud. As the reading went on, some of the students stopped using Relay Reader altogether (ID01, ID02, ID03, ID07, ID11) and some continued reading but stopped producing scorable oral responses at a certain point and did not resume doing so (students ID09 and ID04 in Table 1). Six students persevered in reading out loud through the end of the book. We note that although students could change how much they listen to and how much they read as often as they wanted, only four of the students in Table 1 ever touched the reading settings to move it away from the default of “200-word narrator 150-word student.”⁷ This explains the relative consistency of the number of turns per chapter across

readers. In this respect, reader ID05 stands out as having recorded more turns than the other students in the early chapters of the book (this is the student who made multiple recordings of some of the passages).

Table 1. Number of Scored Oral Reading Turns, per Chapter, per Student

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
01	6	5	5	7	7	3^a													
02	0	0	0																
03	6	6	6	5	7	6	5	7	7	4	4^a								
04	7	5	5	5	7	1	7	8	1	4^a	2	1	0	0	0	0	0	0	0
05	10	6	6	10	9	7	7	7	8	4	4	4	5	8	4	6	6	4	8^a
06	7	5	6	6	7	6	6	7	8	4	4	4	5	8	4	6	6	4	8^a
07	7	5	6	8	7^a	0													
08	7	5	6	7	7	6	6	6	8	4	4	4	5	8	4	6	6	4	8^a
09	7	4^a	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	6	3	6	7	7	6	5	6	8	4	4	4	4	8	4	6	6	4	8^a
11	6	5	6	7	7	5	6	5	3^a	2									
14	7	5	6	7	7	6	6	7	8	4	3	4	5	8	4	6	6	4	8^a
17	7	5	6	7	7	6	6	7	8	3	4	4	5	8	4	6	6	4	8^a

Note. Final chapter for all students apart from ID02 is shaded gray and marked in bold. Zeros indicate that the student read the chapter (based on the event logs) but did not produce any scorable oral reading. An empty cell indicates that the student quit using the app before reaching the chapter.

^a Final chapter for all students apart from ID02 is shaded gray and marked in bold.

To further examine the longitudinal trends, we compared, for every student, average wcpm score across all scored turns in Chapter 1 to the average wcpm score across all scored turns in the furthest chapter for the given student that had at least three scored reading turns.⁸ The final chapters used for the calculation for every student are shown in bold in Table 1. Additionally, we utilized the model from Beigman Klebanov et al. (2020) that evaluated a number of variables for their ability to explain variance in fluency measurements in U.S. upper elementary school children who were reading *Harry Potter and the Sorcerer's Stone* on Relay Reader. Specifically, Beigman Klebanov et al. introduced a “chapter” variable to examine any longitudinal trends in fluency measurements and found it to be a significant predictor after controlling for random passage and reader effects and fixed effects known to impact fluency measurements: the complexity of the passage measured using TextEvaluator® (TE) and the

constraints it imposes on oral production of speech captured by the words per minute rate (wcpm) of a text-to-speech (TTS) model. The final model from Beigman Klebanov et al. (2020) is shown below in lmer syntax (see original paper for further explanation):

$$\text{wcpm} \sim (1|\text{passage}) + (\text{chapter}|\text{student}) + \text{grade} + \text{TE} + \text{TTS} + \text{chapter}$$

After fitting the model to their data, Beigman Klebanov et al. (2020) reported coefficients for the dependent variables as shown in Table 2.

Table 2. Model Fit Results on Data From 35 Upper Elementary Students Reading Harry Potter

Independent variables	Standardized coefficient (SE)
Grade 5	-0.733 (8.909)
TTS	4.841*** (0.932)
TE	-3.036*** (0.911)
Chapter	1.278*** (0.257)
Constant	99.959*** (7.510)

Note. TTS = text to speech; TE = TextEvaluator. Dependent variable is wcpm (word count per minute). $N = 1,556$. Source: Beigman Klebanov et al. (2020).

*** $p < 0.001$.

We fit the same model without the Grade variable to a set of 787 observations collected from the 12 Chilean high school students; descriptive statistics of the data and data preparation steps are described in Appendix B. The resulting model is shown in Table 3. Grade was omitted since the teacher picked students for their strong English language skill relatively to peers across high school grades, so we did not expect grade-level trends in this group of students. In addition, given that we had only 12 students from four grade levels, it was unlikely that we would have been able to faithfully capture any emergent grade-level trends. Indeed, adding Grade to the model did not significantly improve the model’s fit (log likelihoods of -3282.2 vs. -3280.5, respectively, $p = 0.06$). In contrast, omitting the “chapter” variable resulted in a model with an inferior fit (log likelihoods of -3282.2 vs -3292.2, $p < 0.001$). We also observed that removing the possibility of different slopes—rates of growth—across students by using (1 | student) instead of (chapter | student) in the model resulted in a model with a significantly

inferior fit (likelihoods of -3282.2 vs. -3289.3 , $p < 0.001$), therefore we kept the different slopes in the model.

Table 3. Model Fit Results Using the Current Data

Independent variables	Standardized coefficient (<i>SE</i>)	
TTS	6.139***	(0.860)
TE	-3.804***	(0.861)
Chapter	1.235	(0.805)
Constant	116.054***	(7.734)

Note. TTS = text to speech; TE = TextEvaluator. Dependent variable wcpm (word count per minute). $N = 1,556$. Source: Beigman Klebanov *et al.* (2020).

*** $p < 0.001$.

We observe that the general modeling results are in alignment with the prior work, notwithstanding a different book and a different population of readers: TTS and TE are significant predictors, with a positive coefficient for the former and a negative for the latter, and the chapter variable enters the model with a positive coefficient, suggesting a trend toward higher wcpm in later chapters, although the trend does not reach significance in our case ($p = 0.172$). This general alignment with prior results indicates good parsimony of the model. Although the lack of significance for the chapter variable does not support a claim of improvement as students read, we believe that the fact that the coefficient is not negative can provide additional evidence of student *engagement*: Even though the reading was done at home and there was no incentive to make an effort to read fluently, students continued reading conscientiously throughout. Growth modeling will have to be revisited with a larger sample of students in future work.

Another analysis we conducted as a means of exploring the appropriateness of the reading measure was to inspect the oral reading fluency of the readers. Using the mean of the average Chapter 1 scores used in the previous analysis, we found that, on average, students read correctly 120 words per minute—an estimate that coincides with the fluency of a fourth-grade student in a U.S. elementary school, according to published oral reading fluency norms (Hasbrouck & Tindal, 2017).

With regard to the second research question—that is, the students’ and teachers’ perceptions of the technology-mediated ER practice and the application—we observed

generally positive perceptions vis-à-vis the reading experience with Relay Reader. As shown in Table 4, five students responded to the pre- and postsurveys. Additionally, 11 students responded to the in-app survey questions after reading the first chapter and eight students answered the survey questions featured in the app after the final chapter. Overall, all students liked reading in the Relay Reader application on a tablet or mobile device and, on average, perceived reading in the application as easier than reading on their own. They reported no challenges in using the application and generally agreed that Relay Reader helped them improve their English and, in particular, their reading skills in English.

In terms of technology-enhanced features in the application that aim to facilitate the reading process, students shared their opinions about the narrator (see Table 5). All students agreed that the narrator spoke clearly, did not read too fast, and had a positive effect on their ability to comprehend spoken English. However, there seemed to be a slight decrease in the participants' enthusiasm for listening to the narrator. After the first chapter, students still seemed to largely appreciate the alternating process of reading and listening. By contrast, the postsurvey responses and responses provided after Chapter 19 indicated that two out of eight students who responded came to dislike having to listen to the narrator and to alternate between reading and listening toward the end of the study; these two students also adjusted the settings of the activity to minimize listening time and maximize reading time (see endnote 4). The same two students also highlighted speed and optionality of the narrator as a desirable feature to improve the app, pointing out that they did not like the narrator and would prefer the option to turn off the narrator and the option to control the speed of the narrator. Additionally, four out of five students reported that they occasionally followed along in the text when the narrator was reading, but only two indicated that they always followed along.

Table 4. Overall Perceptions of the Technology-Mediated Reading Experience

Perception	I think this app is helping me become a better reader.	Reading felt easier with the app than on my own.	I liked reading with this app.	The app was difficult to use.	I liked the experience of reading on the tablet.	I would prefer to read from a regular (non-electronic) book.	I think this app helped me improve my English.	I would like to read more books in class using the app.	I feel confident reading a full-length book in English.			
<i>n</i>	11	8	11	8	5	5	5	5	5	5	5	
Chapter	Ch 1	Ch 19	Ch 1	Ch 19						Pre-reading	Post-reading	
Average	3.36	3.63	3.1	3.38	3.4	1.4	3.4	2.6	3.2	3.0	3.0	3.2
Strongly agree	4	5	4	4	1	0	2	0	1	0	2	2
Agree	7	3	4	3	4	0	3	3	4	5	1	2
Disagree	0	0	3	1	0	2	0	2	0	0	2	1
Strongly disagree	0	0	0	0	0	3	0	0	0	0	0	0

Note. Responses are on a scale of 1 to 4 (1 = *strongly disagree*; 4 = *strongly agree*). Between chapter surveys 1 and 19, all 13 participants are represented.

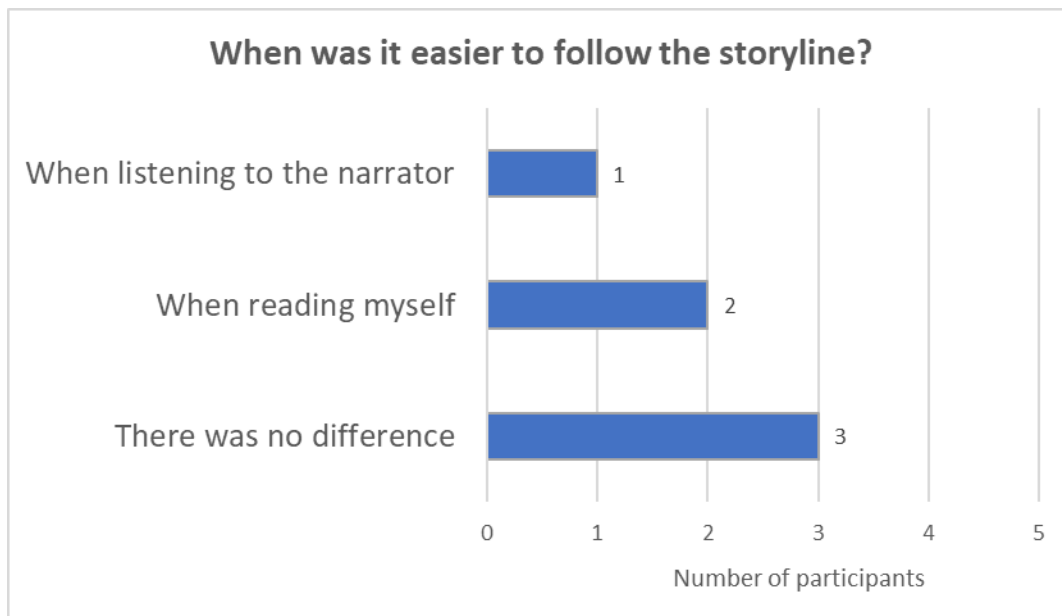
Table 5. In-App Features: Narrator

Perception	I liked alternating between reading and listening.		I liked listening to the narrator.		The narrator spoke clearly.	The narrator read too fast.	I would prefer to read by myself, without the narrator.	I would prefer to read my turns silently, not aloud.	Listening to the narrator improved my ability to understand spoken English.
<i>n</i>	11	8	11	8	5	5	5	5	5
Chapter	Ch 1	Ch 19	Ch 1	Ch 19					
Average	3.64	3	3.73	3.13	3.6	1.4	3.0	2.8	3.4
Strongly agree	7	2	9	4	3	0	2	1	2
Agree	4	4	1	2	2	0	1	2	3
Disagree	0	2	1	1	0	2	2	2	0
Strongly disagree	0	0	0	1	0	3	0	0	0

Note. Responses are on a scale of 1 to 4 (1 = *strongly disagree*; 4 = *strongly agree*).

Also, in response to the question “When it was easier to follow the storyline?” (see Figure 4), three out of five participants emphasized that they noticed no difference in comprehension regardless of whether they themselves were reading or listening to the narrator. Two students indicated that they found it easier to follow the story when they were reading themselves, whereas one person indicated that listening to the narrator facilitated the comprehension process for them.

Figure 4. Perceived Difficulty of Following the Storyline in Human Versus Self-Read Mode



Finally, a preference to read silently was expressed by three of the students in the postreading survey (see Table 5) and clearly enacted by those who read silently either from the beginning or from some point onward (see Table 1). In addition, four out of the five students in the postsurvey said that they never listened to their own readings. However, ID04, who read aloud consistently through Chapter 10 and then moved to silent reading, mentioned being able to read out loud in response to the question what she liked most about the app, which suggests that she was okay doing this initially but preferred silent reading later. In addition, some readers did take advantage of the self-teaching opportunity by perusing their own narrations: One student indicated in an open-ended response in the postsurvey that he found it helpful to

listen to himself because he “can hear [his] errors,” and another student not only listened to her own recordings but also re-recorded herself multiple times.

Regarding the comprehension questions (Table 6), participants generally seemed to like the questions. Two participants highlighted the comprehension question as the feature in the application that they liked the most. One student, for example, noted that “the questions, and how they show the book” were helpful, while another one emphasized the usefulness of the questions insofar as they kept his motivation and focus. He noted that what he “liked the most were the questions as they made [him] unable to get away from reading and in a way forced [him] to be attentive to listen.”

Table 6. In-App Features: Comprehension Questions

Perception	I liked answering questions.	I liked being asked questions as I read.	The questions were distracting.	The questions were difficult.	The questions helped me understand the story.	
<i>n</i>	11	8	5	5	5	
Chapter	Ch 1	Ch 19				
Average	3.45	3.38	3.6	1.6	1.6	2.6
Strongly agree	5	3	3	0	0	0
Agree	6	5	2	0	0	3
Disagree	0	0	0	3	3	2
Strongly disagree	0	0	0	2	2	0

Note. Responses are on a scale of 1 to 4 (1 = *strongly disagree*; 4 = *strongly agree*).

In line with the students, the teacher reported a generally positive view of the Relay Reader application and provided a more detailed account of the observations he made when his students were reading in the application. He shared information about the context of implementation, features in the app, the biweekly reports, and the book (*The Adventures of Pinocchio*) in the Relay Reader app at the time of data collection. Additionally, he shared observations about individual students who participated in the ER practice and ideas for the improvement of the Relay Reader application to provide support for English learning.

With regard to the context, he provided more details about how he implemented the ER practice in the first interview. He noted that due to the global pandemic at the time of data collection, the school was following a fully remote teaching schedule and teachers were struggling with low attendance in online classes. As a means of keeping learners engaged in the schoolwork delivered via the computer, the school followed a policy to not give students homework. Thus, teachers were not allowed to require students to do any work at home. Therefore, he implemented the ER experience as a voluntary, extracurricular activity. In particular, he stated that

we're [i.e., teachers] not allowed to work on any project that involves using home time. So, now they are using like the app to practice like it's kind of hidden way of making them read, right? So they're playing, using an app, they are trying . . . they actually studying so that's what we're doing. We did it as a side project right, it's not part of a program or the class. They're doing it as a side project. (Interview I)

Learners were told that they would receive extra credit if they participated and finished reading the book in the application. Grades were not tied to correct responses to comprehension questions or any other reading related features. Even though it was implemented as an "at home practice" (Interview I), the teacher made himself available via a Teams chat so participating students could reach out if they had questions or were seeking support. However, at the time of the first interview, that is, 2 weeks into the ER practice, none of the students had reached out to him about the ER practice.

In terms of the in-app features, he indicated in the survey responses that he found the narrations very useful because they allowed learners to listen to an English speaker and follow along in the text during the turns read by the narrator. He regarded it as very helpful for students to improve their English listening and reading skills and in particular their English pronunciation. Moreover, he found the reading comprehension questions useful to maintain students' focus and commented positively on the self-recording functionality insofar as it provided an opportunity for students to listen to themselves reading. In particular, he added that

they [i.e., the students] love listening to themselves, but they are too shy to actually do something like this. If we invited them, like [. . .] 10 of them, I think judging by the names, not even five will turn the cameras on. But now, that they have a chance of doing it by themselves, nobody looking, that's fine, they like that. (Interview 1)

As apparent in this quote, he highlighted that at this age students are very self-conscious, so the application allowed them to practice reading and listening without anyone watching. In short, it provided a safe environment for practice. The teacher commented in reference to the biweekly report he received during the ER practice that

it's surprising to me that some of the kids are usually not doing anything and are actually participating this. There's another kid [. . .] he's always saying he's not interested in anything like everything's boring, everything's lame. But now I can see he's very interested in Pinocchio. So, that's pretty cool to know. [. . .]

I will say kids are pretty different when they are alone and they don't have to pretend or they don't have to be worried about the image or whatever they are projecting to the others and they're just going to be themselves they can just use English freely because nobody is judging them or nobody's laughing at them or whatever. So, that thing is very interesting to me and I can see there that they are more interested in English, in learning, in reading than what they show in class. That's interesting to me.

In addition to being able to see the reading progress in the biweekly reports as indicated by the previous quote, the teacher reported in the survey responses that he considered the following pieces of information provided in the report as very useful: list of students with no reading data, reading comprehension statistics including total number of questions answered and percentage correct, daily app use time, and students' current chapter location. Moreover, he regarded the following measures as particularly useful and worthy of including in future teacher reports: cumulative results on all variables from the beginning of the book, how well a student is reading, including intonation and pronunciation, as well as indications of specific words that students pronounced incorrectly.

A particular aspect he regarded as not useful was the reporting of the students' level of oral reading fluency in reference to the U.S. school system. Instead, he noted that it would be more helpful to have the reading levels presented in relation to the CEFR. For example, when asked during the first interview about whether the comparison to the U.S. school system made sense to him and was useful, he responded:

Not really. It's interesting to know about it, but we're not using that in any way. No. But I think it'll be better if you do it like with the European Framework reference because that will be universal. Like here . . . I mean I'm a teacher but I know like almost nothing about the school's system you have there. Like if I have to guess, I don't know how . . . you told me like a 6th grade kid will read. Okay, I have an idea for that words, but I'm not really sure, like I can't compare. Like in my mind sounds good like okay that's not, like it's not a baby. It's a kid, but I don't know which level they really are or anything so it's interesting to know about it, but it's not really informative to me. (Interview 1)

With regard to the book, *The Adventures of Pinocchio*, he commented in the survey that the title may be misleading insofar as it "immediately suggests that it is a children's book." However, in the first interview he added that after starting to read, the students were positively surprised. The teacher recalled that "when they saw like the story, oh this is Pinocchio. Really, I'm 15 years old, [. . .] then they tell like it was a different story like it's not . . . this is not Disney. I was, oh, no it's not. So that was it." Accordingly, they had initially assumed that the book featured a children's story, but soon realized that it was different from the more widely known Disney adaptation.

When asked about individual learners and whether their performance deviated from his observations in class, he reported that he was positively surprised by some students' performance. He highlighted that he "didn't think all of them were like good readers" (Interview 1). For instance, in reference to one student he recalled the following:

I was checking the reports and he is one of the guys like he's been there like for 5 hours or something and he never reads. Never ever. The app that called his attention or something but he never reads. When we have to read books on class, he just doesn't want to do it. And he says, oh, I'll watch a movie. That's boring, I'm not going to do it

but now he's reading and he's spending a lot of time there. I don't know if it's related to the situation like he's in his room doing nothing or he's actually interested but that was surprising. Yeah, that was surprising to me. (Interview 1)

Along these lines, the teacher voiced his surprise when being told in the postinterview that eight of the students finished reading the book, a number that exceeded his expectation that a maximum of one or two students would complete the reading. He added that when he wrapped up the ER practice, some students approached him that they wanted to continue reading in the application. For instance, he referenced one student who said "I want to keep using the app. If they have more books, I want to read the other books too" (Interview 2).

In reference to potential improvements of the Relay Reader application, the teacher made three distinct suggestions across both interviews. First, he highlighted the value of a teacher management system that would not only provide teachers with the students' reading history and statistics, but that would also allow them to listen to recordings of students' reading. Second, he emphasized that feedback that was directly provided by the application to students may help them improve language abilities. For example, he noted that L1-specific pronunciation feedback may be helpful for learners, providing the following example for L1 Spanish learners of English:

Here in Chile, we tend to use an E before the S. Like we don't have any words starting with S here. Like S, a consonant like S . . . I will say space will be *espacio*. So, it will be *especial*. So, they will tend to say *es-special*. So, that kind of thing. (Interview 1)

Lastly, he suggested to implement a functionality in the application that blocks push notifications from other apps in order to avoid students getting distracted if they read on their smartphones.

Finally, in the survey the teacher had indicated that he had most frequently used the following reading materials in his classroom: textbooks, newspaper/magazines, and practical/functional materials (e.g., emails, manuals, etc.). Only sometimes had he included literary or fictional texts. In the postreading interview, the teacher noted that going forward, he wanted to continue using the application to provide his students with ER practice. He highlighted that he

would like to use [the application] as a side grade like reading and I will give them like a book . . . probably I'll do the same thing we do with the real books. I'll give them a book for the whole semester, they can read it in their own times, in their own terms, whenever they want to read. They got a full semester to read one book. I think that's fair enough. One book for a full semester, 5 months, one book that's okay and I will give them like a grade at the end of it. That's the way I would like to use it. (Interview 2)

Discussion

Research Question 1: To what extent do students engage with the Relay Reader application?

Our analyses of the survey responses, the teacher interviews, and the reading samples suggest good engagement and focus among readers, even though the incentives were not tied to oral reading or comprehension performance. One factor to explain this finding is students' enjoyment of many aspects of the activity (turn-taking, listening, questions, the story) and their perception of its utility for improving their English—even after spending hours on the activity. Another factor could be the flexibility of implementation that allowed students substantial agency (without, however, losing accountability, since the teacher could monitor their progress); indeed, schedules varied widely, between a student who logged 16 daily reading sessions lasting from 4 minutes to 1 hour 29 minutes to a student who finished the book in three sittings, one of them more than 4 hours long. An additional contributing factor could be the tool's availability across different technological platforms (web, iOS, Android). The teacher believed that the ability to read on their personal devices contributed to the students' acceptance of the activity. Based on the logs, the students used all the platforms that were offered: Seven of the students predominantly used iOS, four students used Android, and two students used the web application. Two of the students made use of two different platforms.

Although the teacher noted that the activity achieved the intended instructional goal for the students who participated in the ER practice, some students did not give it a chance insofar as they did not even open the application. A potential factor mentioned by the teacher was the perceived “childishness” of the story, possibly based on familiarity with the classic Disney movie or picture-book adaptations. Interestingly, all five students who read the book in Spanish were among the 13 readers; perhaps they knew the story to be quite different from the movie.

Notwithstanding the eventual good engagement for many readers, the teacher recommended horror, mystery, and romance as likely to appeal to the target age group. Along similar lines, students themselves recommended to include additional texts in the Relay Reader app such as fiction like “Percy Jackson, Harry Potter, Grimm brothers books” as well as literary classics like “The Odyssey, The Iliad, The Little Prince,” suggesting that *The Adventures of Pinocchio* was generally in line with the students’ interests for reading in English.

In terms of appropriateness, the strong self-reported reader engagement, a trend toward more fluent reading at the end vis-à-vis the beginning of the oral reading activity, and the teacher’s satisfaction suggest that the activity overall was appropriate for the skill level of the group. The group generally looked quite similar to what one might see in a group of U.S. fourth graders in the middle of the school year in terms of fluency: nine out of 12 student readers fell in the middle 50% of the range, one was a little below 10th percentile, and two were slightly above the 75th percentile of U.S. fourth-grade midyear norms (Hasbrouck & Tindal, 2017). These results align with prior research carried out with fourth- and fifth-grade students in the United States (Beigman Klebanov *et al.*, 2019, 2020; Licalde *et al.*, 2022), suggesting that the adolescent EFL learners in the current study generally fell within the target skill range for the Relay Reader application. In other words, the level of oral reading fluency of the high school EFL learners in this study (approximately B1 on the CEFR) was commensurate with that of students who had participated in prior research conducted with the Relay Reader application in the United States.

Research Question 2: What are the teacher’s and students’ perceptions of the technology-mediated ER practice, in particular, with regard to feasibility of implementation, complexity of the reading passage, and perceived student motivation?

As a whole, the teacher’s and students’ perceptions of the technology-enhanced reading experience, in general, and the application, in particular, were very positive with all participants commenting on the ease to use the application as well as its perceived potential to improve English skills (Table 2). Most in-app features, such as the comprehension questions or the control of the length of reading turns, were perceived as enriching the reading process. By contrast, although all five students who completed the final survey indicated that listening to

the narrator improved their ability to understand spoken English, their enjoyment of the narration and the alternation between reading and listening varied: Out of the eight students who finished the book, two disliked the experience and expressed a wish for the narration to be an optional feature, while two others singled out the alternation between listening and reading as the best feature of the app. The opportunities offered by the app, where students could listen to themselves and re-record their turns after listening to the narrator's performance, were not utilized by most of the readers; nevertheless, a few learners took advantage of these features and found them to be useful. For example, it is interesting to note that only one student engaged in the activity in a way that suggests focus on explicit and deliberate practice of skill, in that she listened to her own recordings, listened to the narrator reading her passages, and re-recorded some of these passages. The rest of the students engaged in such activities very rarely or not at all, instead treating the activity as a book reading activity, namely, focusing on moving ahead with the story. In sum, these mixed results and different usage patterns may be indicative of different preferences in terms of learning. In line with providing multiple means of representation (CAST, 2018), offering learners a variety of ways to interact with the material may be most useful. For example, the option to adjust the extent of reading versus listening enabled the two readers who disliked the listening to reduce the listening to reading ratio to 1:3 and perhaps allowed them to persevere with the activity. Therefore, giving users the option to control the narrator's speed, offering narrators with accents from a range of English varieties, and/or making the narrator an optional in-app feature altogether may be potential directions for development to accommodate the diverse learning needs of readers. Similarly, additional oral reading practices based on performance indices and/or feedback could be included as an optional feature that may provide welcome practice to those readers interested in honing their EFL skills.

Additionally, the technological capability of Relay Reader offered a new way of implementing ER practice in L2 instruction. Given that classroom time tends to be limited, the teacher noted that he wanted to continue offering voluntary ER outside of the classroom using the application. Technological enhancements, such as the narrator, could fulfil support activities that are often carried out by the teacher during the English lesson such as the reading

aloud before students read on their own (Tanaka & Stapleton, 2007). Thus, the teacher could utilize the classroom time to build on the learner's reading activity—a step that would be facilitated further if the teacher were able to track students' progress and review their reading activities by means of a learning management system.

Finally, in line with previous studies (e.g., Day & Bamford, 1998), the technology-mediated nature of the ER activity also seemed to be motivating for students as suggested by the fact that the number of students completing the book had exceeded the teacher's expectations and that some students even asked to continue reading in the application after the intervention.

Concluding Remarks

Overall, this study was a first step to explore a technology-mediated ER experience in the context of a fully remote young learner EFL classroom. It showed that students seemed motivated to engage with the Relay Reader app and that the app with the book *The Adventures of Pinocchio* was indeed appropriate for the group of adolescent EFL learners. It also provided preliminary insight into a way in which ER practice utilizing the Relay Reader application could be implemented into formal EFL instruction without impeding on the limited classroom time that tends to be typical of EFL lessons.

As a final note, there are a number of limitations that need to be recognized as well as directions for future research both in terms of ER practice in general and with the Relay Reader application in particular. First, due to the COVID-19 pandemic, the sample size was very small with only a single school being able to participate in this study. Thus, future research will need to be conducted with a larger pool of participants across multiple schools and learning contexts in order to provide generalizable results. Additionally, this study mainly relied on self-reported user perception data. Therefore, any perceived benefits about the ability of the application to increase learners' English skills will need to be investigated empirically before any claims can be made. For instance, future research could investigate the relationship between students' oral reading fluency (turn scores/wcpm scores) and their reading comprehension as a way to further explore the generally positive perceptions toward the app as well as the design features of it. Finally, we were able to provide preliminary insights into how teachers may use the Relay

Reader app for ER practice over the course of a semester; however, future research will need to document systematically this approach, potential in-class supports, and especially the role technology plays in the entire ER practice.

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*, 23(3), 285–303.
<https://www.jstor.org/stable/748043>
- Beigman Klebanov, B., & Loukina, A. (2021). Exploiting structured error to improve automated scoring of oral reading fluency. In I. Roll, D. McNamara, S. Sosnosvsky, R. Luckin, & V. Dimitrova (Eds.), *Lecture notes in computer science: Artificial intelligence in education, AIED 2021* (Vol. 12749, pp. 76–81). Springer. https://doi.org/10.1007/978-3-030-78270-2_13
- Beigman Klebanov, B., Loukina, A., Lockwood, J. R., Licalalde, V. R. T., Sabatini, J., Madnani, N., Gyawali, B., Wang, Z., & Lentini, J. (2020). Detecting learning in noisy data: The case of oral reading fluency. In C. Rensing & H. Drachsler (Chairs), *LAK20: Proceedings of the 10th International Conference on Learning Analytics & Knowledge* (pp. 490–495). ACM. <https://doi.org/10.1145/3375462.3375490>
- 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 (Chairs), *LAK19: Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 106–110). ACM. <https://doi.org/10.1145/3303772.3303833>
- Bernstein, J., Cheng, J., Balog, J., & Downey, R. (2020). Artificial intelligence for scoring oral reading fluency. In H. Jiao & R. W. Lissitz (Eds.), *Application of artificial intelligence to assessment* (pp. 51–76). Information Age.
- CAST. (2018). *Universal design for learning guidelines* (Version 2.2).
<http://udlguidelines.cast.org>
- Cobb, T., Greaves, C., & Horst, M. (2001). Can the rate of lexical acquisition from reading be increased? An experiment in reading French with a suite of on-line resources. In P. Raymond & C. Cornaire (Eds.), *Regards sur la didactique des langues secondes* (pp. 133–153). Éditions Logique.

- Cox, K. E., & Guthrie, J. T. (2001). Motivational and cognitive contributions to students' amount of reading. *Contemporary Educational Psychology, 26*(1), 116–131.
<https://doi.org/10.1006/ceps.1999.1044>
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge University Press.
- Elley, W. B. (1992). *How in the world do students read? IEA Study of Reading Literacy*. International Association for the Evaluation of Educational Achievement.
<https://eric.ed.gov/?id=ED360613>
- Elley, W. B. (2001). Literacy in the present world: Realities and possibilities. In L. Verhoeven & C. E. Snow (Eds.), *Literacy and motivation: Reading engagement in individuals and groups* (pp. 225–242). Erlbaum.
- Elley, W. B., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly, 19*(1), 53–67. <https://doi.org/10.2307/747337>
- ETS. (2020). *Handbook for the TOEFL Junior® Tests*. https://www.language-testing-service.de/wp-content/uploads/2018/06/TOEFL-YSS-Junior-Tests_Examinee-Handbook_ENG_07122020.pdf
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139150484>
- Guthrie, J. T., Schafer, W. D., & Huang, C.-W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *Journal of Educational Research, 94*(3), 145–162.
<https://doi.org/10.1080/00220670109599912>
- Hafiz, F. M., & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal, 43*(1), 4–13. <https://doi.org/10.1093/elt/43.1.4>
- Hasbrouck, J., & Tindal, G. (2017). *An update to compiled ORF norms* (Technical Report No. 1702). University of Oregon. <https://files.eric.ed.gov/fulltext/ED594994.pdf>
- Hasselgreen, A., & Caudwell, G. (2016). *Assessing the language of young learners*. Equinox Publishing.

- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review*, 61(3), 355–382.
<https://doi.org/10.3138/cmlr.61.3.355>
- Hsueh-chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
<https://doi.org/10.26686/wgtn.12560354.v1>
- Licalalde, 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(1), 16–36. <https://doi.org/10.1037/edu0000532>
- 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. In G. Kubin & Z Kačič (Chairs), *Proceedings Interspeech 2019* (pp. 21–25). International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2019-2889>
- Madnani, N., Beigman Klebanov, B., Loukina, A., Gyawali, B., Lange, P., Sabatini, J., & Flor, M. (2019). My turn to read: An interleaved e-book reading tool for developing and struggling readers. In M. R. Costa-jussa & E. Alfonseca (Eds.), *Proceedings of the 57th annual meeting of the Association for Computational Linguistics: System demonstrations* (pp. 141–146). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/P19-3024>
- Mehigan, G. (2020). Effects of fluency oriented instruction on motivation for reading of struggling readers. *Education Sciences*, 10(3), 1–23.
<https://doi.org/10.3390/educsci10030056>
- Misra, A., Loukina, A., Beigman Klebanov, B., Gyawali, B., & Zechner, K. (2021). A good start is half the battle won: Unsupervised pre-training for low resource children’s speech recognition for an interactive reading companion. In I. Roll, D. McNamara, S. Sosnosvsky, R. Luckin, & V. Dimitrova (Eds.), *Lecture notes in computer science: Artificial intelligence in education, AIED 2022* (Vol. 12748, pp. 306–317) Springer.
https://doi.org/10.1007/978-3-030-78292-4_25

- Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, 49(1), 6–37. <https://doi.org/10.1002/tesq.157>
- Nation, I. S. P. (2008). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. Routledge. <https://doi.org/10.4324/9780203891643>
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Renandya, W. A., & Jacobs, G. M. (2002). Extensive reading: Why aren't we all doing it? In J. C. Richards & W. A. Reynada (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 295–302). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667190.041>
- Schmitt, N. (2008). Instructed second language vocabulary. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177/1362168808089921>
- Siegel, A. (2014). What should we talk about? The authenticity of textbook topics. *ELT Journal*, 68(4), 363–375. <https://doi.org/10.1093/elt/ccu012>
- Tanaka, H., & Stapleton, P. (2007). Increasing reading input in Japanese high school EFL classrooms: An empirical study exploring the efficacy of extensive reading. *The Reading Matrix*, 7(1), 115–131. http://www.readingmatrix.com/articles/tanaka_stapleton/article.pdf
- Tsang, W.-K. (1996). Comparing the effects of reading and writing on writing performance. *Applied Linguistics*, 17(2), 210–233.
- Vogels, E. A., Gelles-Watnick, R., & Massarat, N. (August 10, 2022). *Teens, social media and technology 2022*. Pew Research Center. <https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/>

Appendix A. Relay Reader Teacher Protocols

Relay Reader: Teacher Early-Implementation Interview Protocol

- 1. Can you tell us how you implemented the reading experience?**
 - a. Is it part of a classroom or an extracurricular activity?
 - b. Weekly tasks? Or a final deadline?
 - i. Students are at different stages in reading. Why?
- 2. Could you describe a typical reading session?**
- 3. Did any students experience any technical problems using the app so far?**
- 4. Did any of the students need other help when using the app (help with unfamiliar words, help with understanding the text, answering questions etc.)?**
- 5. Did your students make any comments about the app that you can share?**
 - a. Ex: Student reading for 4 hours ({name of student})
 - b. Ex: {name of student} & {name of student} over 1 hour sessions
- 6. Did you implement additional activities to support your students' reading experience?**
- 7. Did you have a chance to look at the reports we sent with student reading indices?**
 - a. What did you think about them?
 - b. What did you find particularly helpful?
 - c. Was anything confusing?
 - d. Is there any additional information you would like to have included in the report?
- 8. Would you want to have additional reading materials for those students who will most likely be done reading the book soon?**
 - a. Explain the beta version with additional texts

Relay Reader: Teacher Post-Implementation Interview Protocol

1. **Do you think your students liked using the app? Why or why not?**
2. **In your opinion, how important was the experience of reading on the tablet as opposed to a physical book for student engagement?**
3. **How useful was the app for the following purposes?
Did you notice a difference between kids who have stronger or weaker English proficiency to start with?**

[INTERVIEWER: Refer to completed survey; prompt for information about:

- *Improvement in general English language proficiency*
- *Improvement in English listening comprehension*
- *Improvement in English pronunciation*
- *Improvement in reading comprehension in English*
- *Improvement in reading confidence (students felt they could read a book in English)*
- *Improvement in reading independence (students felt they could read a book in English by themselves)*
- *Positive reading experience (students felt the experience was relatively easygoing)*
- *New English learning experience (listening to the narration and reading aloud in turns)*
- *Increased reading stamina or persistence in English (students were comfortable reading multiple times per week)*
- *Social cohesion in the class (through a shared book reading experience)]*

4. **How useful were the following features of the app for engaging your students with a book in English?**

[INTERVIEWER: Refer to completed survey; prompt for information about:

- Narration—listening to the narrator
- Narration—following along with the narrator in the text
- Listening to themselves read
- Listening to the narrator read students' turns
- Reading comprehension questions
- Being able to re-record themselves
- Being able to control the length of their own and narrator turns
- The specific book used in the app (Pinocchio)
- The specific narrator used in the app (Mark Smith)
- App tutorial
- Report screen in the app (under "Reading History")]

5. **Did any of your students need additional incentives to keep on reading? If so, how did you motivate them (e.g., extra credit, no homework)?**
6. **Did you use any activities to support your students' reading experience? If so, which ones did you use (e.g., explicit instruction in key reading strategies; vocabulary activities; class discussion of the story; story-related writing exercises; etc.)? How did you implement them?**
7. **What other genres of books (please include specific titles, if possible) do you think your students would be interested in reading through this app?**
8. **Would you use this app in your future classes? Why or why not?**
[INTERVIEWER: Probe for practical or logistical issues]
9. **How useful were each of the following types of information (provided in the weekly reports sent by the ETS team) for you as a teacher in supporting and monitoring the reading activity?**
[INTERVIEWER: Refer to completed survey; prompt for information about:
 - Turn duration summary (too long, too short)
 - List of students with no reading data
 - Reading comprehension statistics (total questions, % correct)
 - Daily app use time
 - Student's current chapter location/bookmark
 - Graphic of time spent reading aloud compared to other productive/unproductive activities]
10. **How did you use the feedback?**
11. **How useful and informative would additional types of feedback about students' reading using this app be? How would you use this feedback?**
[INTERVIEWER: Refer to completed survey; prompt for information about:
 - A more detailed breakdown of the students' activities (for example, time spent on the "reading history" tab)
 - More real-time feedback (for example, being able to see a report of skipped turns and turn duration in real-time)
 - Cumulative results on all variables from the beginning of the book, not just for the current week
 - How well a student is reading, including fluency, intonation, and pronunciation (even if these data are not available immediately)

- Specific words that students pronounced incorrectly (even if this data is not available immediately)
 - Specific words that a student indicated needing help with (assuming the app gave them an opportunity to indicate these)
 - Report of rewardable achievements, such as a certain number of consecutive comprehension questions correct]
- 12. In addition to all the types of feedback covered above, are there any other types of feedback you would like to see in a report? Please describe.**
- 13. Do you teach in any other settings? Do you think the app would work in that setting?**
- 14. Overall, what do you think of this app?**
- 15. Do you have any other comments that we could use to improve this app?**

Appendix B. Data Preparation for the Analysis of Longitudinal Trends

To support the longitudinal analysis of students' fluency as they read different passages from the book, we removed recordings that constituted re-readings of a previously recorded reading turn (since a re-reading could be faster due to familiarity with the specific passage) and all passages that were read by five people or fewer, in order to create a dataset where passage and student variance could be separated. This resulted in removal of 134 recordings of 107 unique passages.

The dataset for the longitudinal analysis consists of 787 recordings of 108 unique passages, with the average of 7.3 recordings per passage (min = 6, max = 11). The passages span all the 19 chapters, with at least 4 passages per chapter, except for chapter 13 that has 2 passages. There are 5.7 passages per chapter, on average.

ch001	7
ch002	5
ch003	6
ch004	7
ch005	7
ch006	6
ch007	6
ch008	6
ch009	8
ch0010	4
ch0011	4
ch0012	4
ch0013	2
ch0014	8
ch0015	4
ch0016	6
ch0017	6
ch0018	4
ch0019	8

Finally, we verified that there is no significant linear relationship between the prosody (TTS) and the complexity (TE) variables and chapter, to ascertain that neither of these variables

inadvertently code for passage position in the book; the correlations are Pearson's $r = -0.043$ and $r = -0.037$, respectively ($n = 108$).

Notes

- ¹ Note that traditionally young learners have been defined as “children from about 5 years to 17 years” that have been further subdivided into three age groups of “younger children (5/6 to 8/9 years), older children (8/9 to 12/13 years) and teenagers (12/13 to 17 years)” (Hasselgreen & Caudwell, 2016, p. 1).
- ² Note that, in this paper, we use the terms *online*, *remote*, and *distance learning environment* synonymously to indicate that students are not learning EFL face-to-face settings, but instead are receiving instruction via technology-mediated formats (for a description of the study context see below).
- ³ A newer version of the app, released in September 2021, features a library of books.
- ⁴ Note that the content and set up of the research form is equivalent to an operational TOEFL Junior Standard test.
- ⁵ The approximate length of a reading turn in words is set in the reading settings in the app and can be anything between 70 and 200 for each narrator and reader. The actual number of words in each turn is defined by the paragraph structure of the text, given that there are no transitions between the narrator and reader midparagraph. For example, if the target length of the reader is 150 words, paragraphs will be added to the reader turn until the 150-word threshold is crossed. If removing the latest paragraph would result in a turn length that is closer to 150 than retaining that paragraph, the paragraph will be removed. The turns are calculated and allocated dynamically; reading settings can be changed by the reader at any time with the change taking effect immediately in the next turn.
- ⁶ Turns with clear oral reading are operationalized as turns where an automated speech recognition system was able to recognize at least 70% of the text in the passage that the student was supposed to read in the relevant turn. Note that a low proportion could correspond to silent reading, nonreading, or actual oral reading where a recording was of insufficient acoustic quality.

- ⁷ One of the students (ID04 in Table 1) changed the settings 13 times. All students started at 200 narrator and 150 student turns. ID04 explored longer turns (200 each), short reader turns with the long narrator turns, then the reverse, until eventually settling on the longest possible reader turn (200 words) and the shortest possible narrator turn (70 words), in Chapter 7, which the student kept through the end of the book. Reader ID09 changed to this setting after the very first chapter; ID02's final change before quitting the app was also to this setting. These same three students were the ones who either did not read aloud or stopped doing so; the combination of minimizing narration and avoiding oral reading suggests that these students might have had, or developed, a preference for silent independent reading in English.
- ⁸ Taking into account performance on multiple passages is a common practice in assessing oral reading fluency (Bernstein et al., 2020).