Received: 2 October 2020 | Revised: 18 June 2021 | Accepted: 18 July 2021

DOI: 10.1111/jcal.12592

ARTICLE

Journal of Computer Assisted Learning WILEY

Assessing the effectiveness of a game-based phonics intervention for first and second grade English language learners in India: A randomized controlled trial

Priyanka Patel[1] | Minna Torppa[1] | Mikko Aro[2] | Ulla Richardson[3] | Heikki Lyytinen[4]

[1]Department of Teacher Education, Faculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland

[2]Department of Education, Faculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland

[3]Centre for Applied Linguistics, University of Jyväskylä, Jyväskylä, Finland

[4]Department of Psychology, University of Jyväskylä, Jyväskylä, Finland

Correspondence
Priyanka Patel, Department of Teacher Education, Faculty of Education and Psychology, University of Jyväskylä, B325, Alvar Aallon katu 9, Jyväskylä 40600, Finland.
Email: priyanka.v.patel@jyu.fi

Abstract

Background: In 2018, it was found that only a quarter of Grade 3 children in India were reading at grade level. A growing demand for English education has further limited children's literacy achievement. Despite a strong evidence base in favour of using systematic phonics for building English literacy skills, many teachers in India continue to use rote-methods of literacy instruction.

Objectives: We aimed to examine the efficacy of GraphoLearn (GL) English Rime, a computer-assisted reading intervention, in improving the foundational literacy skills of 1st and 2nd grade students who were attending an English medium school in India.

Methods: A total of 136 students across 6 classrooms were randomly allocated to play either GL or a control math game over a 5-week intervention period. Students were pre- and post-tested on various English literacy skills using tasks built into the GL software as well as through oral and paper-based tasks.

Results and Conclusions: Students who played GL showed significantly greater and faster development on in-game measures of letter-sound knowledge, rime unit recognition, and word recognition as compared to students who did not play GL. In addition, GL resulted in greater effects on these measures for students with stronger English literacy skills prior to the start of the intervention. No differences were found between groups on the oral and paper-based tasks.

Implications: GL was able to quickly and effectively teach critical sub-skills for reading. However, a lack of effects on the out-of-game measures opens the door for further discussion on the successful implementation of such interventions.

KEYWORDS

educational technology, English language learners, GraphoLearn, India, intervention, phonics

1 | INTRODUCTION

Today, the importance of the ability to read is largely uncontested. Failure to acquire functional literacy skills is known to have adverse and long-lasting individual and societal effects. Nevertheless, a global learning crisis persists with many children, particularly those in low-income countries, lacking basic reading skills (The World Bank, 2018). India is home to over 200 million school-going children. However, poor quality schools have resulted in consistently low learning levels across the country (Banerjee et al., 2007). In 2018, for instance, it was found that only 27% of Grade 3 students in India were reading at grade-level (ASER, 2018).

Researchers have identified a multitude of factors believed to be contributing to such statistics. A lack of quality early childhood educational opportunities means that many children enter primary school lacking foundational skills identified as critical for school-readiness (ASER, 2019). Upon entry into primary school, children are often faced with textbooks and curricula which fail to take into consideration their existing learning levels (see Pritchett & Beatty, 2012). In addition, a highly multilingual society with ever increasing pressures to learn English means that many children attend school in a foreign language (Graddol, 2010). Lofty expectations on behalf of policy makers, school administration, and society, coupled with outdated pedagogical practices, ultimately results in a situation where many children fall behind early on and are never able to catch up (MHRD, 2019).

Computer-assisted learning programs, particularly in the form of serious games, are growing in popularity around the world to help support student learning. A notably appealing feature of such games is that they allow for a greater personalization of learning; content and levels of difficulty can be designed so that it adapts to the needs of the player allowing for more effective practice. This type of technology-supported personalized learning has been found to show promise for improving learning outcomes, particularly in low and middle-income countries (see Major & Francis, 2020 for review). India has a booming educational technology sector, but there are still many barriers to the effective implementation of technology in schools (Byker, 2014). Furthermore, a lack of efficacy studies in the context of India have made it difficult to identify which types of technologies work best, when, and with whom (Miglani & Burch, 2019). In the present study we aimed to test the efficacy of GraphoLearn English Rime (GL), an empirically validated technology-based reading intervention with first and second grade English language learners (ELLs), in India.

1.1 | Early reading skills and ELLs

Fundamental to learning to read in an alphabetic language, such as English, is the acquisition of the alphabetic principle, an understanding of how symbols of the written language (graphemes) represent sounds in the spoken language (phonemes; Castles et al., 2018). Letter knowledge, which includes the understanding of both letter-sounds and letter-names, and phonological awareness have been identified as foundational skills underlying this alphabetic insight (Byrne, 1998; Byrne & Fielding-Barnsley, 1989). Many studies have found letter-name knowledge to be a powerful predictor of reading achievement, particularly in the early grades (see Foulin, 2005 for review). It acts as a scaffold for learning letter-sounds, a critical skill which aids in the ability to decode unfamiliar words (Huang et al., 2014). In addition, phonological awareness skills, in particular phonemic awareness, the ability to identify and manipulate sounds (phonemes) in spoken words, is essential in helping early readers further connect speech to print (Ehri & Roberts, 2006). In a meta-analysis conducted by Melby-Lervåg et al. (2012), phonemic awareness was found to be one of the strongest correlates of differences in children's word reading ability.

Research examining ELLs and L1 English speakers has indicated that English word reading development occurs along a similar trajectory in both groups. Muter and Diethelm (2001) studied Kindergarten children from multilingual backgrounds who were attending school in English. While language measures such as English vocabulary differentiated L1 English and ELLs, phonological awareness measures did not. Furthermore, English phonological awareness, particularly phonological segmentation ability, and letter knowledge were found to be significant predictors of English reading skills, both concurrently and a year later for both groups.

Similar findings have also emerged when researchers have studied children from Indian L1 backgrounds. Chiappe and Siegel (1999) studied Grade 1 children from Punjabi-speaking homes who were studying in English while living in Canada. They too found that there was little difference in the performance profiles of L1 English and L1 Punjabi children who were learning to read in English. No significant differences were found between the two groups on measures of English phonological processing and word recognition. In addition, both groups relied on using letter-sound correspondences when reading unfamiliar words, and in both groups, those who struggled were less skilled at applying this knowledge. While studies on the reading development of ELLs in India are limited, a recent examination of the reading profiles of children across Grades 1–5 who were learning English as a second literacy language in India revealed phoneme awareness to be a strong predictor of English reading skills (Nakamura & De Hoop, 2014). Phonological awareness and letter knowledge play a significant role in English reading ability, for both L1 English and ELL children. Pedagogically, this opens the door to the possibility of using methods to support these skills for both native and non-native English speakers, as they may be equally valid for both groups of children (Muter & Diethelm, 2001).

1.2 | The role of phonics

Longitudinal studies examining the developmental trajectories of reading and spelling have shown that ELLs who may be behind are able to catch up if they receive adequate literacy exposure and instruction (Lesaux et al., 2007). Systematic phonics, in which children are explicitly taught letter-sound (grapheme-phoneme) correspondences, has repeatedly been identified as the most effective approach to helping children acquire knowledge of the alphabetic principle and to use this knowledge in decoding unknown words (Castles et al., 2018; Ehri, 2003). A notable review on the efficacy of systematic phonics instruction, conducted by the National Reading Panel in the United States, showed a moderate effect ($d = 0.41$) of phonics-based instruction as compared to various other forms of instruction, such as whole-word (Ehri, 2003). In addition, the effect of phonics-based instruction was found to be stronger ($d = 0.55$) when instruction was received early rather than after Grade 1 ($d = 0.27$). The positive effects of systematic phonics instruction have similarly been extended to English language learners (Stuart, 1999).

The small number of studies examining the efficacy of phonics-based methods for teaching English to children in India have thus far been highly promising. Dixon et al. (2011) tested the efficacy of Jolly Phonics intervention as compared to regular classroom instruction, with children attending an English-medium, low-income private school in Hyderabad, India. Within each school, Grade 1 students were given Jolly Phonics intervention for an hour a day, 5 days a week, over a period of 6 months as compared to business-as-usual instruction. Results showed a statistically significant difference between the groups, with the Jolly Phonics group performing better on tasks of reading, spelling, and sounding out letters and words, leading authors to conclude that phonics-based methods are effective for teaching English to children in India, even to those who are first-generation ELLs.

A second study examined the effects of phonics-based English instruction when used with economically disadvantaged Grade 5 children who were attending Kannada-medium school, but who had been learning English as an additional language starting from Grade 3 (Nishanimut et al., 2013). The intervention group received a conventional synthetic phonics intervention for an hour per day while the control group continued to learn English through rote learning of their textbooks. After a 5-week intervention period, those who received phonics-based instruction performed significantly better than the comparison group on measures of letter naming, word reading, non-word reading, and multiple measures of grapho-phonological awareness. The study provided further evidence for the efficacy of phonics instruction, even for those students who are not studying in English-medium school but who are learning English as an additional language. While studies for the use of phonics in the Indian classroom have shown promise, phonics has yet to find a permanent place in the Indian classroom.

1.3 | English literacy instruction in India

Given India's multicultural landscape, it is not uncommon for individuals to use a variety of languages in everyday life. English plays a critical role in India's education system as one of the two official languages of the country (NCERT, 2005), and also plays an important societal role as a language of opportunity (Annamalai, 2004). Consequently, parents from all backgrounds are increasingly choosing to educate their children in schools where English is the primary medium of instruction. As of 2015, 29 million children are believed to be enrolled in English medium schools across India (Nagarajan, 2015) and it is estimated that 25% of children in these schools are first generation ELLs (Graddol, 2010).

It has become clear that while English medium schools are highly desirable, enrolment has not equated to learning. The 2016 Annual Status of Education Report, which specifically examined English literacy in India, showed that 53% of Grade 1 students were unable to identify capital letters, and only 15% of Grade 2 and 20% of Grade 3 students were able to read simple words such as 'fan' (ASER, 2016). A 2012 study on the teaching of English in public primary schools highlighted the many factors contributing to low reading achievement in India, including the predominate focus on rote-reading and copy writing (Dutta & Bala, 2012).

In most Indian public schools, English word reading is taught in a similar fashion to what Rayner et al. (2001) describe as the ABC method (also known as the alphabet-spelling method; see Gupta, 2014), in which children are taught letter names which are then used to spell words (e.g. 'double-you-ay-ell-ell is wall'). Common words are taught through sight and most reading in the classroom is done aloud by the teacher (Dutta & Bala, 2012). While phonics-based instruction has been recommended, there is limited evidence of sustained implementation. Gupta (2014) conducted a study observing phonics instruction in two schools in South India where teachers themselves had learned through the alphabet-spelling method. It was seen that when teachers intended to use phonics, instruction on letter-sounds halted after one sound for each of the 26 letters was taught and teachers continued to articulate individual letter-sounds rather than teaching children how to blend together sounds to decode words (e.g., /k/ /l/ /o/ /ə/ /d/ is cloud; Gupta, 2014). In a study examining factors which influence second-language English reading in India it was found that Grade 1 students in economically disadvantaged areas struggled with English decoding and failed to acquire reading skills even by Grade 5; a finding which was attributed to the lack of phonics-based instruction (Shenoy et al., 2020).

1.4 | GraphoLearn

GraphoLearn (GL; formerly known as GraphoGame), is a globally studied computer-assisted reading intervention which uses systematic phonics to train the connections between spoken and written language (Lyytinen et al., 2009; Lyytinen et al., 2021; Richardson & Lyytinen, 2014). Numerous studies have provided evidence for the explicit training of phonological skills and letter-sound linkages in helping poor readers develop understanding of the alphabetic principle (see Hatcher et al., 1994, 2004), even for ELLs (Stuart, 1999). GL promotes these skills using adaptive technology, allowing for individualized practice. Implemented in more than 20 countries across various languages, there is a growing evidence base for the use of GL around the globe (Lyytinen et al., 2021; McTigue et al., 2019).

In this study we used GL English Rime, a version of GL which uses systematic rhyme family groupings, where a small number of individual letter-sound (grapheme-phoneme) correspondences are introduced,

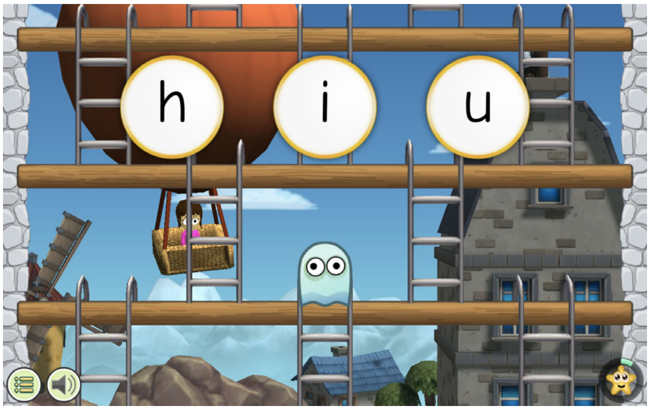

FIGURE 1 Example of a letter-sound game screen from GL English Rime. Players hear a speech-sound and are required to select the correct written form out of the options provided. The number of options increases as players advance within the game [Colour figure can be viewed at wileyonlinelibrary.com]

after which they are combined to form larger and more consistent orthographic rime units, and finally words (see Kyle et al., 2013). The game consists of 25 play streams and 7 assessment streams, each with multiple levels, in which the content is organized so that the largest rhyme families with the most consistent orthographic rime spellings are introduced first. Within each level, players hear a speech sound which they are then required to match with the correct written unit from the multiple options presented on the screen (see Figure 1). Phonological awareness skills are further trained through rhyme awareness tasks and spelling skills are supported through word formation tasks in which players drag letter tiles into the correct order to spell a given word. If players make an incorrect selection, they are provided with feedback which guides them to the correct selection. Players are required to achieve 80% accuracy across the levels within a given stream to unlock the next stream. To aid in motivation, rewards are provided throughout the game in the form of stars and coins. For research purposes, data from the game is automatically saved to an external server when players quit their play session so long as the device has an active internet connection.

A review of findings across 28 GL studies highlighted that the effectiveness of GL varies across languages and educational contexts (McTigue et al., 2019). In this study, we specifically evaluated the effectiveness of the English Rime version of GL. The first study utilizing GL English Rime was conducted by Kyle et al. (2013) in which GL English Rime was compared to both, a phoneme-based version of GL which focuses on only instructing individual letter-sound correspondences (vs. coupled with rime units) and an untreated control group. Both game versions were found to be effective in comparison to the control group, and no significant differences were found in the effect sizes of the gains between the two game versions. When the GL Rime group was compared to the untreated controls, large effects were found on the gain scores for tasks of word reading, non-word reading, and spelling. Bhide et al. (2013) conducted a study in which the effects of GL Rime were compared to a musical intervention. The GL group showed large effects on decoding and spelling, however, there were no significant differences between the groups indicating that both interventions benefited struggling readers.

Based on these two small-scale pilots, a large-scale randomized controlled trial study of GL Rime was conducted (Worth et al., 2018) with almost 400 Grade 2 students across the UK. Teachers found GL Rime easy to implement, engaging, and motivating, however, there was no evidence of improved reading outcomes over business-as-usual instruction. Due to the large sample size, there was great variation in how long children played GL and how far they progressed in the game. A reanalysis of this sample was conducted by Ahmed et al. (2020), in which they specifically examined only those children who reached above the group mean in play progress, or what they refer to as the 'top half' of players. When these children were compared to the full sample of control children, those in the GL group showed significantly higher gains and it was concluded that GL is more effective than business-as-usual in developing English phonics (Ahmed et al., 2020). Wilson et al. (2021) recently re-examined the same data set to better understand the types of children who best respond to GL intervention as indicated by game progress. They found that phonological skills and executive functioning skills were the strongest unique predictors of game progress. Interestingly, vocabulary was not a significant predictor indicating that even those with limited English vocabulary are able to benefit from playing GL English Rime.

Findings from these studies are highly informative given that GL has been found to be just as effective as business-as usual literacy instruction, at least in the UK (Worth et al., 2018). When it works, technology-led instruction can benefit children's learning (see Banerjee et al., 2007), often in a time and cost-effective manner (Muralidharan et al., 2019). Often, however, technology-led solutions developed in high-income contexts are 'copy-pasted' into more challenging environments such as those in low-income countries where resources are strained and knowledge is limited, and as a result, they may not work as effectively (Trucano, 2014). Therefore, it is critical that efficacy studies of educational technologies are conducted in various environments to help us untangle what works, for whom, and when (Trucano, 2005).

Currently, there is limited evidence on the efficacy of GL English Rime in places such as the Global South. One study which examined GL English Rime intervention as compared to phoneme and word-level interventions in Singaporean schools found that all three interventions led to increased reading outcomes but there were no significant differences between the three interventions (O'Brien et al., 2019). Unfortunately, however, this study did not contain a no-intervention control group. An extensive amount of GL research has also been done in sub-Saharan Africa where children are learning to read in a multilingual context and are faced with many of the same challenges as students in India. While these studies have not utilized GL English Rime, results from studies implementing other versions of GL in Africa have been highly promising (see Lyytinen et al., 2019 for review). We believe that it is worthwhile to extend the existing evidence base around the efficacy of GL English to include India, a country housing a large portion of the world's student population who are failing to acquire literacy skills.

1.5 | The present study

In this study, we aimed to examine whether GL English Rime could be used as an effective intervention to improve the foundational literacy skills of children learning to read English in India. The present study is an extension of a former small-scale pilot study, the first of its kind to examine GL in India. Grade 3 ELLs in an English-medium public school in Ahmedabad, India were randomly allocated to play either GL English Rime or a control math game as a supplement during school (Patel et al., 2018). Results indicated that those who received GL intervention showed significant gains in letter-sound knowledge after an 8-week intervention period. While these results were promising, the authors found no evidence of a transfer of learning to paper-based tasks of reading and spelling.

This pilot study had some notable limitations, which we believe may help to explain the pattern of results. Along with a small sample size, the study was conducted with Grade 3 children. Meta-analyses have shown a greater overall effect of phonics interventions when used with younger children (i.e. Kindergarten-Grade 1; $d = 0.55$) versus older children (i.e. Grade 2–6; $d = 0.27$; Ehri, 2003). Thus, we aimed to replicate the study with a younger sample of students. We also included a wider range of oral and paper-based assessments and selected assessments which were specifically developed and validated for use on children in India. Our research questions were as follows:

1. Do children who play GL significantly outperform children who do not on game-based measures of English reading skills?
2. Do children who play GL significantly outperform children who do not on oral and paper-based measures of English reading skills?
3. How does progress within GL relate to students' pre-test, post-test, and gain scores? Are there differences in effectiveness based on this relation?

2 | MATERIALS AND METHODS

2.1 | Participants

The data reported is from 136 students across three Grade 1 classrooms and three Grade 2 classrooms from one public school in Delhi, India. The students were an average of 6.2 years old (range = 5–7 years) at the start of the study which began approximately 6 months into the school year. The primary medium of instruction was English but students learned Hindi, the regional language in Delhi, as an additional language for one period per day. Student demographics, for each condition, including age, gender, grade level, and language used at home are described in Table 1.

Prior to the start of the study, the students' parents were invited to the school and taken through the consent form which was provided to them in both English and Hindi to ensure they were fully informed. In total, the parents of all 143 students consented and all participated in the study. However, data from seven students has been excluded from the analysis due to dropping out of the study prior to post-testing or having GL data which failed to save to the server.

TABLE 1 Demographics

	GL ($n = 69$)		Control ($n = 67$)	
	n	%	*n*	%
Gender				
Male	35	51	34	51
Female	34	49	33	49
Grade				
1	33	48	31	46
2	36	52	36	54
Home Language				
Hindi	57	83	55	82
Hindi + English	9	13	9	14
Hindi + Other	3	4	1	2
Other	0	0	1	2

2.2 | Procedure

A matched pairs randomized design was used in which randomization was done within classrooms. Students within each classroom were matched on age and gender, and then randomly allocated to either the GL group ($n = 69$) or the control group ($n = 67$) which played a math game. Teachers were fully informed of the study but were asked to continue teaching their lessons as usual. Prior to the start of the intervention, students were pretested in a group format as well as individually. The individual sessions lasted approximately 30 minutes and all testing was conducted by the primary researcher along with trained research assistants who also facilitated the intervention. After the intervention was complete, post-testing was done using the same measures and by the same facilitators.

For the intervention sessions, 25 smartphones were set up in a spare classroom in the school and students were brought in class-by-class, 5 days a week, for 20 minute sessions during their regular school hours. The students in the intervention group played GL English Rime. Each student had a pre-created avatar which was labelled with their name to ensure that they played under the same profile for the entire duration of the intervention. The students in the control group played a math game called ‘Math Kids- Add, Subtract, Count, and Learn’, which was selected from the Google Play store. ‘Math Kids’ consists of mini games to practice basic counting and comparison skills, as well as basic arithmetic operations; skills the class teachers confirmed students were learning according to their curriculum. Like the GL group, each student in the control group had a profile labelled with their name under which they played. Although there was no pre-set order which the game required players to follow, the research team ensured that children were playing a different, often more difficult, level each week to maintain interest and motivation.

Although the game had very limited visual and auditory language input, the game language was changed to Hindi to ensure that any language exposure was not occurring in English. The purpose of the math game was to ensure that both groups of children were exposed to technology-based games and spent an equivalent amount of time away from classroom instruction.

Fidelity to the intervention was controlled by logs sent to the GL server which include days played and time spent playing. Students' attendance in the sessions was also recorded by the research team, and although students played their respective games independently, the primary researcher and research assistants supervised the play sessions. Upon completion of the intervention, there was no significant difference between the GL group ($M = 20.19$, $SD = 2.42$) and the control group ($M = 19.91$, $SD = 2.69$; $t(134) = 0.60$, $p = 0.55$), in the number of play sessions attended.

2.3 | Measures

Both groups were assessed at pre- and post-test using the GL in-game assessment which contains a letter-sounds task, a rime unit task, and a word recognition task (see Table 2 for detailed task descriptions). The children were brought in class-by-class into a spare classroom where the smartphones were set up. The game was introduced to them after which they were instructed to play the assessment levels. In the assessment levels, just as in the game levels, players are presented with an auditory target which they are required to match with a visual target out of the multiple options provided. The in-game assessment contains both trained and untrained items and not all of the trained content is in the assessments. The purpose of the in-game assessment tasks was to assess players' performance on skills explicitly instructed by the game (i.e., recognition) and in a manner similar to that in which they had learnt and practiced those skills (i.e., matching an auditory target to a visual target). The letter-sound knowledge task contains 24 trials and players are exposed to all trials regardless of performance. The rime units task contains 24 trials, and the word recognition task contains 47 trials; however, these tasks discontinue after the player answers more than 50% of the items incorrectly.

TABLE 2 GL in-game task descriptions

Task	Description	Scoring
Letter sounds	Students were presented individual letter sounds auditorily which they had to match with the correct written form out of the multiple options presented to them	Total number of correct responses
Rime units	Rime units (i.e., -ip, -at) were presented auditorily which students were required to match with the correct written form out of the multiple options presented to them	Total number of correct responses
Word recognition	Words were presented auditorily which students were required to match with the correct written form out of the multiple options presented to them	Total number of correct responses

Oral and paper-based tasks were administered at pre- and post-test to give insight into existing literacy skills at pre-test and to determine if there is a transfer of learning to an out-of-game context at post-test. An important consideration for this study was the use of assessments which were designed and validated for use with children in India. Subtests from the English version of the Dyslexia Assessment for Languages of India (DALI; Rao et al., 2021; Singh, 2015), as well as tasks (PhAB) developed by Cherodath and Singh (2015) were used. Students were brought into a quiet room within the school and the tasks were administered one-on-one by the primary researcher along with trained research assistants. The spelling assessment was conducted as a whole class dictation administered either by the primary researcher or the research assistants. Detailed descriptions of the oral and paper-based tasks, scoring criteria, and their reliability are provided in Table 3.

3 | RESULTS

The data was analysed using IBM SPSS Statistics. First the distributions of the raw scores of the in-game assessments and the oral and paper-based assessments were examined for normality. The in-game assessments contained outliers which resulted in slightly skewed distributions. For all three tasks, the scores were winzorized to meet the assumption of normality. Regarding the oral and paper-based measures, scores on the letter name identification ($M = 9.58$, $SD = 1.11$) and letter spelling ($M = 9.42$, $SD = 1.50$) measures were at ceiling at pre-test, consequently resulting in a negatively skewed distribution. Therefore, these two measures were not analysed further. All remaining oral and paper-based measures produced a normal distribution at both time points. In addition, the two English word reading lists were found to be highly correlated at both pre-test ($r = 0.87$) and post-test ($r = 0.89$), therefore, an average of the two scores was used for analysis.

To answer the first research question, the results from the GL in-game assessment tasks at pre- and post-test were explored (see Table 4). Repeated measures ANOVA was used to examine time × group interaction effects, as well as main effects of time and group. A significant time × group interaction effect was found for all three in-game assessment tasks, with the GL group showing significantly higher scores and faster development than the control group. Effect sizes (partial eta squared) of the interactions are

TABLE 3 Out-of-game oral and paper-based task descriptions

Task	Description	Scoring	Cronbach's α
Rhyme identification (DALI)	Two practice sets followed by 12 sets of three words were presented orally and students were required to identify the two words which rhymed	A score of one was given for every correctly identified pair	0.83
Phoneme replacement (DALI)	Two practice words followed by 10 words were presented orally in which students were asked to replace the initial phoneme with a given phoneme	A score of one was given for every correct replacement	0.89
Letter naming (DALI)	Student were asked to name 10-upper case letters presented to them on a sheet	A score of one was given for every correctly named letter	0.76
Semantic fluency (DALI)	Students were given 30 seconds each to name as many objects in two given categories—fruits and vegetables	The total number of correctly named items in each category were counted	–
Verbal fluency (DALI)	Students were given 30 seconds each to name as many words beginning with two given phonemes—/b/ and /m/	The total number of correctly named words in each category were counted	–
Word reading (DALI)	A set of 25 words which were collated from Grade 1 and 2 textbooks and arranged in order of increasing difficulty	A score of one was given for every correctly read item	0.94
Word reading (PhAB)	A set of 20 words collated from Grade 1–3 textbooks	A score of one was given for every correctly read item	0.94
Pseudoword reading (PhAB)	A set of 20 words in which a single letter in a real word was replaced to create a legally pronounceable string	A score of one was given for every correctly read item	0.91
Letter spelling (DALI)	10-item letter name dictation	A score of one was given for every correctly written letter	0.83
Spelling (DALI)	20-item word dictation consisting of words collated from Grade 1 and 2 textbooks, presented in order of increasing difficulty	A score of one was given for every correctly written word	0.90

TABLE 4 Descriptive statistics and group comparisons on GL in-game assessment tasks

		GL *M* (*SD*)	Range	Control *M* (*SD*)	Range	*t*	Group	Time	Interaction	Effect size η_p^2
Letter sounds	Pre	7.93 (2.93)	0–14	7.75 (2.28)	2–14	0.34	61.99***	210.04***	131.89***	0.50
	Post	17.61 (5.95)	4–41	8.97 (3.79)	2–26	11.11***				
Rime units	Pre	1.55 (1.83)	0–10	1.43 (1.69)	0–9	0.50	35.55***	69.43***	43.89***	0.25
	Post	7.26 (6.10)	0–21	2.12 (2.63)	0–15	6.67***				
Word recognition	Pre	3.41 (2.21)	0–8	3.16 (2.50)	0–14	0.74	5.56*	49.34***	7.86**	0.06
	Post	6.93 (6.99)	0–40	4.36 (3.08)	0–14	2.85**				

***$p < 0.001$. **$p < 0.01$. *$p < 0.05$.

reported in Table 4. The criteria as that defined by Cohen (1988) is being used in which $\eta_p^2 \geq 0.01$ is a small effect, $\eta_p^2 \geq 0.06$ is a medium effect, and $\eta_p^2 \geq 0.14$ is a large effect. In line with the significant interaction effects, a large effect was found for both the letter-sounds and rime unit tasks and a medium effect was found for the word recognition task. An independent samples *t*-test was then conducted to examine group differences at pre- and post-test across the three tasks, and results showed no significant group differences at pre-test, indicating equivalent groups. At post-test, significant group differences in favour of the GL group were found for all three tasks.

To answer the second research question, the results from the oral and paper-based tasks at pre- and post-test were explored (see Table 5). Results of a repeated measures ANOVA indicated no significant time × group interaction effects across all of the measures. A significant main effect for time was found, indicating development in both groups across all tasks, however, no significant main effect was found for group. Nevertheless, effect sizes of the interaction

TABLE 5 Descriptive statistics and group comparisons on out-of-game oral and paper-based tasks

		GL *M* (*SD*)	Range	Control *M* (*SD*)	Range	*t*	Group	Time	Interaction	Effect size η_p^2
Rhyme identification	Pre	4.84 (3.73)	0–12	5.31 (3.00)	0–11	−0.82	0.001	62.33***	3.76	0.03
	Post	7.30 (3.62)	0–12	6.81 (3.53)	0–12	0.81				
Phoneme replacement	Pre	4.19 (3.17)	0–10	4.36 (3.42)	0–10	−0.30	0.13	92.68***	2.56	0.02
	Post	6.65 (3.01)	0–10	6.12 (3.40)	0–10	0.97				
Semantic fluency	Pre	6.94 (3.64)	0–15	7.61 (3.54)	0–16	−1.09	0.48	24.27***	1.21	0.01
	Post	8.38 (3.49)	2–16	8.52 (4.09)	0–19	−0.22				
Verbal fluency	Pre	4.38 (3.49)	0–14	4.27 (3.27)	0–11	0.19	0.000	140.34***	0.23	0.002
	Post	6.83 (3.65)	0–17	6.93 (3.08)	0–12	−0.17				
Word reading	Pre	10.85 (6.66)	0–22.5	11.06 (6.36)	0–22	−0.19	0.01	176.36***	0.41	0.003
	Post	13.78 (6.76)	0–22.5	13.72 (6.73)	0–22.5	0.05				
Pseudoword reading	Pre	5.55 (5.38)	0–19	5.52 (5.41)	0–18	0.03	0.004	68.31***	0.01	0.000
	Post	8.54 (6.43)	0–20	8.45 (6.65)	0–20	0.08				
Spelling	Pre	9.43 (5.18)	0–20	9.66 (4.83)	0–19	−0.26	0.01	25.73***	0.36	0.003
	Post	10.74 (5.01)	0–20	10.68 (4.54)	0–18	0.06				

***$p < 0.001$.

TABLE 6 Game progress and assessment score correlations of the GL group

	Pre-test score	Post-test score	Gain score
In-game measures			
Letter sounds	0.56***	0.69***	0.41***
Rime units	0.33**	0.64***	0.58***
Word recognition	0.50***	0.70***	0.54***
Oral and paper-based measures			
Rhyme identification	0.55***	0.73***	0.20
Phoneme replacement	0.62***	0.67***	0.02
Semantic fluency	0.50***	0.64***	0.15
Verbal fluency	0.69***	0.68***	0.03
Word reading	0.79***	0.81***	0.10
Pseudoword reading	0.67***	0.76***	0.32**
Spelling	0.76***	0.82***	0.06

***$p < 0.001$. **$p < 0.01$.

suggested a small effect for rhyme identification, phoneme replacement, and semantic fluency. Results of an independent samples *t*-test showed no significant group differences at pre-test or post-test across the oral and paper-based measures.

In line with our third research question, we aimed to better understand individual differences in learning from the game by further examining the GL group ($n = 69$) and exploring the relationship between progress made by students in the game (i.e., highest stream reached out of 25) and their scores at pre- and post-test, as well as their gain scores. As previously mentioned, GL requires that players reach 80% mastery within a stream before allowing them to move on to the next stream. Therefore, we wanted to see which children were progressing in the game and whether children's progress in the game related to their performance on the assessment tasks. Stream data as recorded by the GL server was used for the analysis. Correlations between the highest stream reached and the pre-test, post-test, and gain scores across all the assessment tasks are reported in Table 6. A significant positive relationship was found between the highest stream reached in the game and students pre- and post-test scores across all measures, indicating that those who made it further in the game not only had higher post-test scores but also higher pre-test scores. A significant positive relationship was also found between the highest stream reached and the gains made for all the in-game measures as well as the pseudoword reading measure, indicating that those who completed more of the game also had greater in-game assessment gains and pseudoword reading gains.

TABLE 7 Descriptive statistics and group comparisons on the GL in-game assessment tasks for the bottom 50% and top 50%

		GL *M* (*SD*)	Control *M* (*SD*)	*t*	Group	Time	Interaction	Effect size η_p^2
Bottom 50%								
Letter sounds	Pre	6.56 (2.68)	6.31 (3.02)	0.36	28.96***	68.39***	48.17***	0.41
	Post	15.33 (7.22)	7.06 (2.61)	6.98***				
Rime units	Pre	0.89 (1.14)	0.80 (0.87)	0.37	12.96**	24.59***	13.74***	0.17
	Post	4.25 (4.38)	1.40 (1.87)	3.86***				
Word recognition	Pre	2.50 (1.83)	2.09 (1.87)	1.23	3.41	19.79***	1.09	0.02
	Post	4.25 (3.06)	3.09 (2.53)	1.75				
Top 50%								
Letter sounds	Pre	9.42 (2.45)	9.31 (2.83)	0.17	92.68***	181.37***	102.96***	0.62
	Post	20.09 (2.45)	11.06 (3.80)	13.79***				
Rime units	Pre	2.27 (2.16)	2.13 (2.08)	0.41	39.66***	57.20***	40.52***	0.39
	Post	10.55 (6.05)	2.91 (3.10)	6.86***				
Word recognition	Pre	4.39 (2.19)	4.34 (2.59)	0.26	5.24*	24.69***	6.87*	0.10
	Post	9.85 (8.76)	5.75 (3.06)	2.65*				

***p < 0.001. **p < 0.01. *p < 0.05.

Given these findings, we further explored whether there were differences in game effectiveness based on children's pre-test levels. To do this, a composite score was calculated of all of the pre-test measures (both GL in-game and oral and paper-based). The full sample was then divided using median splits resulting in two new groups, one with students who had English literacy skills in the top 50% at pre-test ($n = 33$ GL, $n = 32$ control) and one with students who had English literacy skills in the bottom 50% at pre-test ($n = 36$ GL, $n = 35$ control).

Pre- and post-test means, standard deviations, and group comparison results on the GL assessment tasks for the students grouped by pre-test performance are reported in Table 7. Repeated measures ANOVA was once again used to examine time × group interaction effects, as well as main effects of time and group across the GL assessment tasks for both the top 50% and bottom 50% (see Table 7). For the bottom 50%, a significant time × group interaction effect was found for both the letter-sounds and rime unit tasks, and the effect sizes of the interaction were large. On the word recognition task, however, no significant interaction effect for time × group was found. Nevertheless, the effect size of the interaction indicated a small effect. There was a significant main effect of time, but no significant main effect of group indicating that both groups developed over time. For the top 50%, a significant time × group interaction effect was found across all three tasks with the GL group showing significantly higher scores and faster development than the control group. Effect sizes of the interaction were large for both letter-sounds and rime units, and medium for word recognition.

An independent samples *t*-test was then conducted to examine group differences at pre- and post-test across the three in-game tasks. For both the bottom 50% and the top 50%, no significant group differences were found at pre-test between the GL and control groups, thus indicating equivalent groups. At post-test, the bottom 50% had significant group differences in favour of the GL group only for the letter-sounds and rime unit tasks. In the top 50%, however, significant group differences at post-test in favour of the GL group were found for all three tasks. Taken together, these findings indicate a stronger effect of GL for children with better pre-existing English literacy skills prior to the start of the intervention. However, significant time × group interaction effects were not found on the oral and paper-based measures, indicating that even in those children who had better pre-existing literacy skills prior to the start of the intervention, the skills learned in GL did not transfer to the oral and paper-based measures.

4 | DISCUSSION

In this study, we aimed to examine whether GL English Rime, a globally recognized computer-assisted reading intervention, could significantly improve the foundational English literacy skills of Grade 1 and 2 ELLs who were attending an English-medium public school in Delhi, India. At the end of a 5-week intervention period, the GL group made significant improvements, particularly on in-game assessments of letter-sound knowledge, rime unit recognition, and word recognition.

Along with significantly higher post-test scores, children who played GL showed faster development across all three in-game measures as compared to children who did not play GL. These results are meaningful in showing that GL was able to quickly and effectively teach letter-sound correspondences, a critical subskill for English word reading, to young ELLs in India. In addition, children displayed that they were able to use this newly acquired knowledge to recognize larger units, such as orthographic rimes, and even words. Oral and paper-based tasks, used to determine if there was a transfer of skills learned in the game to non-game-based tasks of reading and spelling, indicated that there was no transfer. This finding is in line with previous GL studies across various languages which have found positive effects of GL on reading subskills but have failed to see in-game outcomes translating to out-of-game measures, particularly word-level reading (see McTigue et al., 2019). Nevertheless, we believe that these findings are meaningful given that that the children in this study were ELLs and the intervention was carried out over a short duration.

As previously mentioned, GL is adaptive and requires an 80% passing criterion. Consequently, children progress through the game at their own pace. While this is advantageous in allowing for individualized practice, from an efficacy perspective, there are challenges that emerge as a result of differences in players progression. As in previous studies (Worth et al., 2018), in this study there was great variability in children's game progress. Thus, in an attempt to better understand response to GL intervention, we examined if and how game progress related to children's pre-test, post-test, and gain scores. In general, we saw that the children who had higher pre-test scores (i.e., better pre-existing English literacy skills) were also the ones who were progressing further in the game. On the contrary, those children who had more limited English literacy skill were perhaps unable to meet the 80% criterion as quickly and consequently, made less progress given the limited intervention period. Based on these findings, we divided children into new groups based on their pre-test scores which allowed us to examine whether there were differences in the effectiveness of GL for children with different pre-existing English literacy skills as compared to their matched controls. At least for the GL in-game assessment tasks, there were differences in game effectiveness with those children with better pre-existing English literacy skill seemingly benefitting more. These findings are in line with previous GL English Rime studies which have found that existing phoneme awareness skill is predictive of response to GL (Wilson et al., 2021) and those children who do respond to GL seem to make gains (Ahmed et al., 2020).

These findings essentially demonstrate a Matthew effect, a phenomenon commonly discussed in relation to reading (Stanovich, 1986) but also one which has been discussed in relation to educational technology (Trucano, 2013). Children who have some basic level of competencies to engage with reading and technology, will read and engage more, and as a result, reap the greatest benefits. However, given that GL was designed as a practice tool in which it is assumed that children have some level of prior phonological knowledge (Richardson & Lyytinen, 2014), these results are perhaps not all that surprising. When children play GL in their native language, they are able to use their existing phonological awareness skill as a foundation upon which to build and learn from the game. On the contrary, for children like those in this study who lack foundational skills in English, GL alone is perhaps not maximally beneficial as indicated by a lack of effects on the oral and paper-based measures.

4.1 | Game on or game over?

There is no doubt that educational technology has great potential to enhance learning experiences. Computer-assisted games can aid in building motivation, confidence, and excitement for learning. However, from an educational perspective, the larger goal of using such games is to teach skills and not just to teach the game. In other words, it is essential that learners can transfer the skills learned in a game to out-of-game contexts. While a detailed discussion on the mechanisms of transfer is beyond the scope of this paper, we will discuss three broad elements of this study which may have contributed to the findings. The first is regarding the measures that were used, the second is regarding game-based factors, and the third is regarding the method of implementation.

A recent meta-analysis examining the effects of educational apps across 36 intervention studies emphasized that 'measures matter' (Kim et al., 2021). Researchers found larger treatment effects in studies which used researcher-developed versus standardized measures and in studies which measure constrained skills (e.g., letter-sound knowledge) versus unconstrained skills (e.g., word reading/vocabulary). Other studies have also shown greater transfer for trained items than untrained items (Görgen et al., 2020; Hintikka et al., 2008). In this study, the oral and paper-based tasks used were standardized measures which were highly unconstrained. In addition, they contained none of the trained items from GL and required children to move beyond simple recognition. In line with previous research, the largest effects were found on the GL tasks in which children were assessed only on those items taught in the game and in a manner most like the game. It is possible that the use of oral and paper-based measures which were more closely aligned to the content taught in GL would have resulted in greater effects.

Although GL supports the development of phonological awareness skills, the phonological awareness tasks used in this study, both phoneme substitution and rhyme oddity, are known to be complex and cognitively demanding (Vandervelden & Siegel, 1995; Wagner & Torgesen, 1987), particularly for children who are ELLs (Pufpaff, 2009). Given that the children in this study had only been learning English for a limited time and there was no classroom emphasis on building phonological awareness skills, it is possible that the phonological awareness tasks were simply too difficult for children to master after a limited intervention period. Many of the oral and paper-based tasks were also distant from what children had practiced in GL. Phoneme replacement was not a task which was explicitly practiced in GL and regarding word reading, in GL, children were practicing word recognition and not oral word production as required for the out-of-game measure. It is also important to remember that out of the three in-game measures, the smallest effects were seen on the word recognition task. Therefore, it is possible that children simply did not build up enough in-game word recognition skill to see a transfer of learning to out-of-game word reading.

Other explanations may lie in the design of the game itself. For one, GL's adaptive features means that at the end of an intervention period children will have only completed as much of the game as they were able to master with 80% accuracy. This also means that there will be variation in how much of the game students will have completed prior to post-testing. In this study, students were able to complete an average of 12 streams (out of 25), with students in the bottom 50% group having completed an average of 7 streams and the students in the top 50% group having completed an average of 17 streams. As mentioned in previous studies, significant effects on decoding were found when children were able to play at least 16 streams (Ahmed et al., 2020). It is difficult to determine if, or how, the results of this study would have changed had more students progressed further or reached game completion, but this is something that warrants further study. Another aspect which requires future exploration is whether the adaptive features of GL are in fact beneficial for these children. Vanbecelaere et al. (2020) conducted an intervention study where they included both an adaptive and non-adaptive version of the game-based intervention and found no additional benefit of adaptation. It would be highly informative to replicate this design using GL.

A second, game-based explanation could lie in the in the content of the game. While GL does aim to support the development of children's phonological awareness skills, its emphasis is limited. Ronimus et al. (2020) acknowledged that the versions of GL used in many previous studies do not explicitly train players' blending and segmenting skills. While GL English Rime does contain word formation levels in which children arrange letter tiles to form a word, this is not a primary focus. For children, such as those in our sample who are in classrooms where phonics is not the norm (see Gupta, 2014), a practice tool alone is not enough.

This brings us to our final, but perhaps most important, explanation of the pattern of results which is the method of implementation. In this study, GL was implemented as a supplement to classroom instruction, but teachers were not asked to make any modifications to their literacy instruction in alignment with GL nor was GL modified to align with existing classroom practices. This was done intentionally to mirror the common reality in India where teachers are provided with technology with little to no training on how to effectively integrate that technology into their teaching practices (Central Square Foundation, 2015). The results of our study reinforce that this type of implementation greatly limits the potential benefits of such interventions.

McTigue et al.'s (2019) meta-analysis on GL found that while a lack of transfer to word reading was prevalent in multiple GL studies across various languages, studies which involved high adult interaction produced an average positive effect ($g = 0.48$; see Saine et al., 2011). Research on other educational technologies have similarly shown that the transfer of learning is enhanced when the learner is provided with various forms of guidance and support (see Tobias & Fletcher, 2007 for review). Such guidance is perhaps even more critical to build certain skills, such as phonological awareness, which are difficult to instruct using technology alone given that they have a large oral component. A previous study examining the effects of a computer-based phonological awareness intervention as compared to teacher-led instruction for young readers found that students in the teacher-led-group significantly outperformed those in the computer-based group on tasks of rhyme discrimination, rhyme production, phoneme isolation, and phoneme segmentation (Mitchell & Fox, 2001). Researchers concluded that while computer-administered instruction allowed children to recognize items presented aurally, it did not give them the opportunity to practice pronunciation, and pronunciation may be required to master the skill. We believe that future GL studies with greater teacher involvement could significantly enhance the effects of the intervention.

4.2 | Limitations

While we believe this study was methodologically strong and makes an important contribution to the limited existing efficacy studies in the Indian context, we recognize that there are limitations. Some of the limitations we have mentioned earlier include a short intervention duration limiting the number of children who could play GL to completion and oral and paper-based measures which were difficult and/or distant from the skills learned in the game.

Regarding the measures used, a limitation of the in-game assessment is the lack of reliability information. Unfortunately, it was not possible to retrieve the reliability information for the version of GL used in this study. However, the reliability of the Finnish version of GL has been examined, and findings have shown that computer-based assessments conducted through GL have high reliability (Hautala et al., 2020). Nevertheless, the GL English Rime in-game assessments require further evaluation. In addition, although we included a wide range of literacy skill assessments, we did not have an oral vocabulary measure which limited our ability to control for children's pre-existing English skills, and thus, we recommend future studies ensure that oral vocabulary is assessed.

Regarding the methods, we did have a limited sample size which results in reduced statistical power. We also would like to acknowledge that the use of a composite score to divide children into the top and bottom 50% groups limits variation and looking at the differences between these groups more in detail is required in future research. Finally, given that India is a highly multicultural and diverse country, it is important to acknowledge that the results of this study are perhaps limited in their generalizability. It is extremely important the efficacy of GL be evaluated in different schools across the country to identify where and with whom GL can be the most beneficial.

4.3 | Practical implications

Overall, the findings from this study shed light on the benefits and constraints of using technology to teach critical skills such as early reading. Policy makers and researchers alike are continuously trying to push the limits of technology. We recognize an urgent need, particularly among the world's most vulnerable populations, to try and quickly improve learning levels before more children fall through the cracks in the system. However, while technology is a means to an end, it is not *the* end.

Research specifically examining the role that technology can play in enhancing reading outcomes is consistently indicating that comprehensive methods in which computer and non-computer based instruction is integrated, and where teachers are provided with professional development, result in greater gains in reading (Cheung & Slavin, 2012; Cheung & Slavin, 2013; Kim et al., 2016; McTigue & Uppstad, 2019). While technology-based interventions allow for individualization, integration is essential in ensuring that there is no misalignment between what is learned in the game and what is instructed in the classroom (Muralidharan et al., 2019).

Educational reform in India clearly goes beyond just the introduction of technological games. There is a clear need to work closely with teachers to help them improve the methods that are being used to teach language and literacy, particularly in English. Technologies, such as GL, can be an effective addition when teachers are adequately trained on when, how, and with whom to use them. In fact, previous research has shown GL intervention to be most effective when used by both teachers and students, rather than just one or the other (Jere-Folotiya et al., 2014). Teachers in India could use similar methods to practice their own understanding of English phonics, which could then better support their instruction. To aid in greater transfer, teachers could integrate GL with classroom literacy instruction to help children make those connections between what they are learning in the game and how they can use that knowledge outside of the game. One way of doing this would be to scaffold learning so that students are not just identifying letter-sounds, rime units, and words in GL but also orally producing them with the teacher.

While classroom instruction into which GL is blended is critical, we understand this relationship between classroom pedagogy and technology should be bidirectional. Therefore, future versions of GL could also be modified to better align with the Indian context. Many children in India come from multilingual backgrounds and it is important to think about how their native language, or other languages they are learning in school, can be utilized in their English literacy learning through GL. Multiple studies in the Indian context have shown children's literacy skills in the first language (L1) to be one of the strongest predictors of their English reading skills (Nakamura & De Hoop, 2014; Reddy & Koda, 2013). Although the examination of the effect of children's L1 skills on their English learning was beyond the scope of this paper, it is highly probable that those children with better English skill, and who seemingly benefitted more from GL, also had better Hindi reading skills. To achieve a more detailed understanding of the skill profiles of the children who can benefit the most from GL in multilingual context, future studies should also include a detailed assessment of L1 literacy skills. Some researchers have suggested aligning English phonics instruction with the stage of phonetic development in children's L1 (Dixon et al., 2011), and others have shown this alignment to be effective, at least with non-technology-based interventions (Nishanimut et al., 2013). Studies replicating these findings using technology-based interventions such as GL could be highly influential.

All in all, the results of this study make an important contribution to research on technology-based English literacy interventions in the context of India and opens the door for topics of future research to enhance the body of evidence around the use of educational technologies to improve literacy in such environments.

ACKNOWLEDGEMENTS

We would like to thank the school, principal, teachers, parents, and students for their cooperation and participation in this study. A special thank you to Dr. Nandini Chatterjee Singh (UNESCO MGIEP) for her support throughout the study and for providing us with the oral and paper-based assessment tools used in this study. We would also like to extend a heartfelt thanks to our research assistants (Ms. Mahima Bhalla, Ms. Samriddhi Dhiman, & Ms. Saumya Goel) for their support during data collection, Mr. Shrey Mishra (Gaming Central India) for helping us acquire the hardware used in this study, and Ms. Sandhya Saxena (UNESCO MGIEP) for her support in the translation of various documents used in this study.

CONFLICT OF INTEREST

The authors declare that the research has been conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jcal.12592.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Priyanka Patel https://orcid.org/0000-0002-1181-1066
Minna Torppa https://orcid.org/0000-0003-3834-9892
Mikko Aro https://orcid.org/0000-0002-0545-0591
Ulla Richardson https://orcid.org/0000-0001-9181-5106
Heikki Lyytinen https://orcid.org/0000-0003-2867-1749

REFERENCES

Ahmed, H., Wilson, A. M., Mead, N., Noble, H., Richardson, U., Wolpert, M. A., & Goswami, U. (2020). An evaluation of the efficacy of GraphoGame rime for promoting English phonics knowledge in poor readers. *Frontiers in Psychology*, *5*(132), 1–12. https://doi.org/10.3389/feduc.2020.00132

Annamalai, E. (2004). Medium of power: The question of English education in India. In J. W. Tollefson & A. B. M. Tsui (Eds.), *Medium of instruction policies: Which agenda? Whose agenda* (pp. 177–193). Lawrence Erlbaum Associates.

ASER. (2016). Annual status of education report (rural) 2016. ASER Centre. http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/aser_2016.pdf

ASER. (2018). Annual status of education report (rural) 2018. ASER Centre. http://img.asercentre.org/docs/ASER%202018/Release%20Material/aserreport2018.pdf

ASER. (2019). Annual status of education report (rural) 2019. ASER Centre. http://img.asercentre.org/docs/ASER%202019/ASER2019%20report%20/aserreport2019earlyyearsfinal.pdf

Bhide, A., Power, A., & Goswami, U. (2013). A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, *7*(2), 113–123. https://doi.org/10.1111/mbe.12016

Banerjee, A., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly*

Journal of Economics, *122*(3), 1235–1264. https://doi.org/10.1162/qjec.122.3.1235

Byker, E. J. (2014). ICT in India's elementary schools: The visions and realities. *The International Education Journal: Comparative Perspectives*, *13*(2), 27–40.

Byrne, B. (1998). *The foundation of literacy: The child's acquisitions of the alphabetic principle*. Psychology Press.

Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, *81*(3), 313–321. https://doi.org/10.1037/0022-0663.81.3.313

Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Central Square Foundation. (2015). The EdTech promise: Catalysing quality school education at scale. http://www.centralsquarefoundation.org/wpcontent/uploads/2015/11/The-EdTech-Promise-Catalysing-Quality-SchoolEducation-at-Scale.pdf

Cherodath, S., & Singh, N. C. (2015). The influence of orthographic depth on reading networks in simultaneous biliterate children. *Brain and Language*, *143*, 42–51. https://doi.org/10.1016/j.bandl.2015.02.001

Cheung, A. C., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*, *7*(3), 198–215. https://doi.org/10.1016/j.edurev.2012.05.002

Cheung, A. C., & Slavin, R. E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly*, *48*(3), 277–299. https://doi.org/10.1002/rrq.50

Chiappe, P., & Siegel, L. S. (1999). Phonological awareness and reading acquisition in English-and Punjabi-speaking Canadian children. *Journal of Educational Psychology*, *91*(1), 20–28. https://doi.org/10.1037/0022-0663.91.1.20

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Earlbaum Associates.

Dixon, P., Schagen, I., & Seedhouse, P. (2011). The impact of an intervention on children's reading and spelling ability in low-income schools in India. *School Effectiveness and School Improvement*, *22*(4), 461–482.

Dutta, U., & Bala, N. (2012). Teaching of English at primary level in government schools. NCERT. http://www.ncert.nic.in/departments/nie/del/publication/pdf/English_Primary_level.pdf

Ehri, L. C. (2003). Systematic phonics instruction: Findings of the National Reading Panel. Paper presented at the invitational seminar by the standards and effectiveness unit, DfES, London.

Ehri, L. C., & Roberts, T. (2006). The roots of learning to read and write: Acquisition of letters and phonemic awareness. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 113–131). Guilford Press.

Foulin, J. N. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing*, *18*(2), 129–155. https://doi.org/10.1007/s11145-004-5892-2

Görgen, R., Huemer, S., Schulte-Körne, G., & Moll, K. (2020). Evaluation of a digital game-based reading training for German children with reading disorder. *Computers & Education*, *150*, 103834. https://doi.org/10.1016/j.compedu.2020.103834

Graddol, D. (2010). English next India: The future of English in India. British Council. https://www.britishcouncil.in/sites/default/files/english_next_india_-_david_graddol.pdf

Gupta, R. (2014). Change in teaching practices: Case of phonics instruction in India. *Procedia-Social and Behavioral Sciences*, *116*, 3911–3915. https://doi.org/10.1016/j.sbspro.2014.01.865

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, *65*(1), 41–57. https://doi.org/10.2307/1131364

Hatcher, P. J., Hulme, C., & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry*, *45*(2), 338–358. https://doi.org/10.1111/j.1469-7610.2004.00225.x

Hautala, J., Heikkilä, R., Nieminen, L., Rantanen, V., Latvala, J. M., & Richardson, U. (2020). Identification of reading difficulties by a digital game-based assessment technology. *Journal of Educational Computing Research*, *58*(5), 1003–1028. https://doi.org/10.1177/0735633120905309

Hintikka, S., Landerl, K., Aro, M., & Lyytinen, H. (2008). Training reading fluency: Is it important to practice reading aloud and is generalization possible? *Annals of Dyslexia*, *58*, 59–79. https://doi.org/10.1007/s11881-008-0012-7

Huang, F. L., Tortorelli, L. S., & Invernizzi, M. A. (2014). An investigation of factors associated with letter-sound knowledge at kindergarten entry. *Early Childhood Research Quarterly*, *29*(2), 182–192. https://doi.org/10.1016/j.ecresq.2014.02.001

Jere-Folotiya, J., Chansa-Kabali, T., Munachaka, J. C., Sampa, F., Yalukanda, C., Westerholm, J., Richardson, U., Serpell, R., & Lyytinen, H. (2014). The effect of using a mobile literacy game to improve literacy levels of grade one students in Zambian schools. *Educational Technology Research and Development*, *62*(4), 417–436. https://doi.org/10.1007/s11423-014-9342-9

Kim, J., Gilbert, J., Yu, Q., & Gale, C. (2021). Measures matter: A meta-analysis of the effects of educational apps on preschool to grade 3 children's literacy and math skills. *AERA Open*, *7*, 1–19. https://doi.org/10.1177/23328584211004183

Kim, Y. S. G., Boyle, H. N., Zuilkowski, S. S., & Nakamura, P. (2016). Landscape report on early grade literacy. USAID. https://www.globalreadingnetwork.net/sites/default/files/eddata/LandscapeReport_0.pdf

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG rime and GG phoneme. *Reading Research Quarterly*, *48*(1), 61–76. https://doi.org/10.1002/rrq.038

Lesaux, N. K., Rupp, A. A., & Siegel, L. S. (2007). Growth in reading skills of children from diverse linguistic backgrounds: Findings from a 5-year longitudinal study. *Journal of Educational Psychology*, *99*(4), 821–834. https://doi.org/10.1037/0022-0663.99.4.821

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E., & Richardson, U. (2009). In search of a science-based application: A learning tool for reading acquisition. *Scandinavian Journal of Psychology*, *50*(6), 668–675. https://doi.org/10.1111/j.1467-9450.2009.00791.x

Lyytinen, H., Ojanen, E., Jere-Folotiya, J., Damaris Ngorosho, S., Sampa, F., February, P., Malasi, F., Munachaka, J., Yalukanda, C., Pugh, K., & Serpell, R. (2019). Challenges associated with reading acquisition in sub-Saharan Africa: Promotion of literacy in multilingual contexts. In N. Spaull & J. Comings (Eds.), *Improving early literacy outcomes: Curriculum, teaching, and assessment, IBE on curriculum, learning, and assessment* (Vol. 4, pp. 119–132). Brill Sense.

Lyytinen, H. J., Semrud-Clikeman, M., Li, H., Pugh, K., & Richardson, U. (2021). Supporting acquisition of spelling skills in different orthographies using an empirically validated digital learning environment. *Frontiers in Psychology*, *12*, 675. https://doi.org/10.3389/fpsyg.2021.566220

Major, L., & Francis, G. A. (2020). Technology-supported personalized learning: Rapid evidence review. *EdTechHub*, 1–36. https://doi.org/10.5281/zenodo.3948175

McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2019). Critically reviewing GraphoGame across the world: Recommendations and cautions for research and implementation of computer-assisted instruction for word-reading acquisition. *Reading Research Quarterly*, *55*(1), 45–73. https://doi.org/10.1002/rrq.256

McTigue, E. M., & Uppstad, P. H. (2019). Getting serious about serious games: Best practices for computer games in reading classrooms. *The Reading Teacher*, *72*(4), 453–461. https://doi.org/10.1002/trtr.1737

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352. https://doi.org/10.1037/a0026744

MHRD. (2019). Draft National Education Policy 2019. https://mhrd.gov.in/sites/upload_files/mhrd/files/Draft_NEP_2019_EN_Revised.pdf

Miglani, N., & Burch, P. (2019). Educational technology in India: The field and teacher's sensemaking. *Contemporary Education Dialogue*, *16*(1), 26–53. https://doi.org/10.1177/0973184918803184

Mitchell, M. J., & Fox, B. J. (2001). The effects of computer software for developing phonological awareness in low-progress readers. *Literacy Research and Instruction*, *40*(4), 315–332. https://doi.org/10.1080/19388070109558353

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, *109*(4), 1426–1460. https://doi.org/10.1257/aer.20171112

Muter, V., & Diethelm, K. (2001). The contribution of phonological skills and letter knowledge to early reading development in a multilingual population. *Language Learning*, *51*(2), 187–219. https://doi.org/10.1111/1467-9922.00153

Nagarajan, R. (2015). Number of children studying in English doubles in 5 years. *The Times of India*. https://timesofindia.indiatimes.com/india/Number-of-children-studying-in-English-doubles-in-5-years/articleshow/49131447.cms

Nakamura, P., & De Hoop, T. (2014). Facilitating reading acquisition in multilingual environments in India (FRAME-India) final report. American Institutes for Research. https://www.air.org/sites/default/files/downloads/report/FRAME_Final%20Report_Final.pdf

NCERT (2005). National Curriculum Framework 2005. National Council of Educational Research and Training. https://ncert.nic.in/pdf/nc-framework/nf2005-english.pdf

Nishanimut, S. P., Johnston, R. S., Joshi, R. M., Thomas, P. J., & Padakannaya, P. (2013). Effect of synthetic phonics instruction on literacy skills in an ESL setting. *Learning and Individual Differences*, *27*, 47–53. https://doi.org/10.1016/j.lindif.2013.06.007

O'Brien, B. A., Habib, M., & Onnis, L. (2019). Technology-based tools for English literacy intervention: Examining intervention grain size and individual differences. *Frontiers in Psychology*, *10*, 2625. https://doi.org/10.3389/fpsyg.2019.02625

Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2018). GraphoLearn India: The effectiveness of a computer-assisted reading intervention in supporting struggling readers of English. *Frontiers in Psychology*, *9*, 1045. https://doi.org/10.3389/fpsyg.2018.01045

Pritchett, L., & Beatty, A. (2012). The negative consequences of overambitious curricula in developing countries. (Center for Global Development Working Paper No. 293, HKS Working Paper No. RWP12-0350).

Pufpaff, L. A. (2009). A developmental continuum of phonological sensitivity skills. *Psychology in the Schools*, *46*(7), 679–691. https://doi.org/10.1002/pits.20407

Rao, C., Sumathi, T. A., Midha, R., Oberoi, G., Kar, B., Khan, M., Vaidya, K., Midya, V., Raman, N., Gajre, M., & Singh, N. C. (2021). Development and standardization of the DALI-DAB (dyslexia assessment for languages of India—Dyslexia assessment battery). *Annals of Dyslexia*, 1–19. https://doi.org/10.1007/s11881-021-00227-z

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*(2), 31–74. https://doi.org/10.1111/1529-1006.00004

Reddy, P. P., & Koda, K. (2013). Orthographic constraints on phonological awareness in biliteracy development. *Writing Systems Research*, *5*(1), 110–130. https://doi.org/10.1080/17586801.2012.748639

Richardson, U., & Lyytinen, H. (2014). The GraphoGame method: The theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Human Technology*, *10*(1), 39–60. https://doi.org/10.17011/ht/urn.201405281859

Ronimus, M., Eklund, K., Westerholm, J., Ketonen, R., & Lyytinen, H. (2020). A mobile game as a support tool for children with severe difficulties in reading and spelling. *Journal of Computer Assisted Learning.*, *36*, 1011–1025. https://doi.org/10.1111/jcal.12456

Saine, N. L., Lerkkanen, M. K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at risk for reading disability. *Child Development*, *82*(3), 1013–1028. https://doi.org/10.1111/j.1467-8624.2011.01580.x

Shenoy, S., Wagner, R. K., & Rao, N. M. (2020). Factors that influence reading acquisition in L2 English for students in Bangalore, India. *Reading and Writing*, *33*, 1809–1838. https://doi.org/10.1007/s11145-020-10047-z

Singh, N.C. (2015). Dyslexia assessment for languages of India (DALI). National Brain Research Centre.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407. https://doi.org/10.1598/RRQ.21.4.1

Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology*, *69*(4), 587–605. https://doi.org/10.1348/000709999157914

The World Bank. (2018). World development report 2018: Learning to realize education's promise. *The World Bank*.

Tobias, S., & Fletcher, J. D. (2007). What research has to say about designing computer games for learning. *Educational Technology*, *47*(5), 20–29. https://www.jstor.org/stable/44429439

Trucano, M. (2005). Knowledge maps: ICT in education. infoDev/World Bank. http://www.infodev.org/en/Publication.8.htm

Trucano, M. (2013). The Matthew effect in educational technology. *The World Bank*. https://blogs.worldbank.org/edutech/matthew-effect-educational-technology

Trucano, M. (2014). Promising uses of technology in education in poor, rural, and isolated communities around the world. *The World Bank*. https://blogs.worldbank.org/edutech/education-technology-poor-rural

Vanbecelaere, S., Van den Berghe, K., Cornillie, F., Sasanguie, D., Reynvoet, B., & Depaepe, F. (2020). The effectiveness of adaptive versus non-adaptive learning with digital educational games. *Journal of Computer Assisted Learning*, *36*(4), 502–513. https://doi.org/10.1111/jcal.12416

Vandervelden, M. C., & Siegel, L. S. (1995). Phonological recoding and phoneme awareness in early literacy: A developmental approach. *Reading Research Quarterly*, *30*, 854–875. https://doi.org/10.2307/748201

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192–212. https://doi.org/10.1037/0033-2909.101.2.192

Wilson, A., Ahmed, H., Mead, N., Noble, H., Richardson, U., Wolpert, M. A., & Goswami, U. (2021). Neurocognitive predictors of response to intervention with GraphoGame rime. *Frontiers in Education*, *6*, 125. https://doi.org/10.3389/feduc.2021.639294

Worth, J., Nelson, J., Harland, J., Bernardinelli, D., & Styles, B. (2018). GraphoGame rime: Evaluation report and executive summary. National Foundation for Educational Research. https://www.nfer.ac.uk/graphogame-rime-evaluation-report-and-executive-summary

How to cite this article: Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2021). Assessing the effectiveness of a game-based phonics intervention for first and second grade English language learners in India: A randomized controlled trial. *Journal of Computer Assisted Learning*, 1–14. https://doi.org/10.1111/jcal.12592