Assessing the effectiveness of GraphoLearn combined with classroom instruction on phonics: a randomised control trial

Deepti Bora

Master's Thesis in Education
Spring Term 2023
Faculty of Education and Psychology
University of Jyväskylä

ABSTRACT

Bora, Deepti. 2023. Assessing the effectiveness of GraphoLearn combined with classroom instruction on phonics: a randomised control trial. Master's Thesis in Education. University of Jyväskylä. Faculty of Education & Psychology.

Proficiency in reading is a gateway for learning. However, the literacy rate in India stands at only 73% today. Many children studying in English-medium public schools often face a learning disadvantage through the schooling years since they struggle in learning to read in English which manifests into a struggle with reading to learn in later years. This disadvantage is more pronounced in children living in poverty. Thus, consistent efforts towards building foundational literacy skills of these children when developmentally appropriate is important to overcome the learning crisis.

The aim of this study was to determine whether GraphoLearn, a computer assisted learning tool, combined with GraphoLearn-aligned classroom instruction can support the foundational English literacy skills of struggling readers in India. Sixty Grade 2 students who were non-native speakers of English and attending an English-medium public school in India participated in the study. Most of the students had at least one year of exposure to English language, however, their literacy skills were not at par with their grade level.

In the intervention, all the students were provided 30-35 minutes of classroom instruction on phonics two to three times a week. In addition, the experimental group was also administered GraphoLearn, and the control group played a math game for 20 minutes five to six days a week at school. While exposure to GraphoLearn led to significant improvement in students' in-game letter-sounds, rime units and word recognition skills, it did not result in transfer to out-of-game measures. However, exposure to the classroom instruction on phonics significantly predicted students' outside-the-game letter-sound knowledge and phoneme blending skills. These results underline the relevance of instruction on phonics and its efficacy in developing literacy skills in struggling readers.

Keywords: GraphoLearn, GraphoLearn-aligned classroom instruction, phonics, foundational literacy

ACKNOWLEDGEMENTS

First and foremost, I would like to thank the state's Department of Education of Education for granting me the permission to conduct this study. I would also like to thank the school principal and staff members for accommodating me as their own and supporting me and this intervention throughout the study period. I would also like to the teachers for their participation and for being flexible and open to the study. The effort they put in to implement the classroom instruction was critical to the success of this intervention. A heartfelt thanks to the 60 students for their participation and being a source of joy and laughter even in most difficult situations. Their patience especially during testing allowed for a seamless data collection process.

Most of all I want to express my sincere gratitude to my supervisor Minna Torppa for her guidance and support. Her observations and feedback during our discussions on the research design, implementation process and data analysis were the driving force behind this study' success. I want to also acknowledge the steady support I received from my friend and colleague Priyanka Patel whose findings on the efficacy of GraphoLearn in India were the backbone of this study. Her expertise in the subject and consistent support gave me more confidence to implement this randomised control trial. Together, both Minna and Priyanka made a dream team for this study.

The implementation of GraphoLearn and its data processing would not have been possible without Mika Halttunen, programmer with the GraphoLearn team at the University of Jyvaskyla, and Ulla Richardson. I would also like to extend my gratitude to both for helping me understand the gaming tool better and providing technical support. The findings of this study are also enabled by the analysis we ran on the data. Towards this, I want thank Maria Psyridou for teaching quantitative data analysis and providing constructive feedback on my work. I would also like to express my gratitude to the dean Anna-Maija for her constant support and confidence in my ability to do this intervention.

I also want to thank my mother who from her experience as a public-school principal provided me with priceless insights into the learning and teaching process in schools. Special mention to my uncle Naresh who provided guidance towards implementing the intervention in a seamless manner. Last, but certainly not the least, I want to thank my friends Bhavani Ramamoorthy and Mohit Singh for standing by my side and being my strength. Thank you all, once again, for bringing this study together.

LIST OF FIGURES AND TABLES

FIGURES

TABLES

CONTENT

1. INTRODUCTION

Literacy is a gateway to lifelong learning. It is fundamental to employability, health, and civic engagement. Yet, despite many advances in education policies and technologies, our struggle to unlock the potential of education continues. Today, 53% of children in low-and middle-income countries cannot read and comprehend a basic text by the age of 10 (World Bank, 2019). This crisis has now been termed as 'learning poverty' by the World Bank to emphasize the importance of achieving basic proficiency in reading. In India, 'learning poverty' continues to ail 55% of children in the late primary years (NCERT, 2021b). As per the National Achievement Survey 2021, the average reading with comprehension level stands at 62% in Grade 3 and 52% in Grade 5 (NCERT, 2021a). This consistent decline in reading level over the schooling years show the perils of the learning crisis as inability to read by age 10 is likely to lead to a lifetime of illiteracy. (NCERT, 2021b; World Bank 2019) To overcome this challenge, India has outlined achievement of universal foundational literacy by 2025 as its highest priority in the National Education Policy (MHRD, 2020).

Literacy development requires provisions such as print-rich classrooms and homes, the use of oral literacy skills in adult-child interactions, and instruction towards building all aspects of literacy in a specific language (Nag, 2013). However, for many children, access to such learning facilities is a struggle as they cannot afford it owing to their poor socio-economic background. Children come from poor families living on only up to $2 a day (PEW, 2021), with little to no provisions for nutrition, clean water, and learning environment at home. Enrolment in public schools does not necessarily provide children access to quality schooling because of challenges such as high teacher absenteeism, deficient teaching methods, and poor classroom environment (Muralidharan & Singh, 2020). Also, studying in a language with no prior exposure or support at home and with inadequate instruction in school often worsens children's chances of education attainment. In this context, education technology (EdTech) is being betted as a tool for equity and social justice (Ojanen et al., 2015).

In many developing countries, EdTech is being employed in schools for improving the access to improving academic skills and overcoming psychological barriers (Escuet2a et al., 2020). EdTech is increasingly used as a lever for enhancing access to quality education from early childhood through the postsecondary years (Escueta et al., 2020) especially for at-risk students (Muralidharan et al., 2019). Amidst various offerings through EdTech, computer-assisted learning (CAL) technology has gained much recognition across as they are intended to overcome the challenge of addressing diverse learning levels within a classroom (Escueta et al., 2020; Rodriguez-Segura, 2022). As a CAL tool, educational games have become popular for their notable features such as scaffolding, adaptive instruction, instant feedback, and content across various levels of difficulty (Plass et al., 2015). However, the implementation and adoption of such tools have often outpaced the evaluation of their efficacy which is needed to identify and understand contexts and people for whom the CAL can be the most beneficial.

This study examines whether GraphoLearn (GL), a computer-assisted learning tool and classroom instruction on phonics together can effectively support the development of foundational English literacy skills among struggling readers. Carried out with students attending an English-medium public school in India, the study also aims to determine the relation between the extent of exposure to the classroom instruction and the development in students. Before presenting the analysis and findings, this thesis has unpacked students' context and using CAL and instruction on phonics for supporting foundational literacy.

1.1 Context

Schools in India follow the three-language policy i.e., children learn to read and write in three languages by the end of secondary school. Typically, two of these three languages include India's official languages – Hindi and English. Towards reading and writing development in the said languages, literacy learning in India today calls for fulfilment of three curricular goals. Laid out in India's National Curriculum Framework for Foundational Stage (NCERT, 2022), these goals are expected to be fulfilled by the time children turn 8 years old. In the first curricular goal, children are expected to build communication skills in their L1 and L2/L3 for daily interactions. The second goal emphasises on fluency in reading and writing in L1. Lastly, the third goal expects children to begin reading and writing in L2/L3 (NCERT, 2022). To support the development of these goals, emphasis has been placed on instructional approaches that use bilingual or multilingual interactive activities for oral language development, providing a print-rich environment, and building decoding skills (NCERT, 2022).

The above three curricular goals emphasise the steep climb that lies ahead of children given their current reading levels (ASER Centre, 2022; NCERT, 2021a). In the Annual Education Survey Report (ASER Centre, 2022), a large-scale, countrywide survey, it is indicated that the percentage of Grade 3 students who can read a Grade 1 and 2 level text is reported to be 30% and 20.5% respectively. Moreover, only 42.8% of Grade

5 and 69.6% of Grade 8 students can read a Grade 2 level text. (ASER Centre, 2022) In their English language ability, only 55.3% of Grade 3 students could read and tell the meaning of simple words such as *sun*, *bus*, and *cat*. At Grade 5 and 8, only 24.5% and 46.7% of students respectively can read simple English sentences - *I like to read.*; *She has many books*. (ASER Centre, 2022) While India is inching towards its goal of universal access for almost 250 million children, clearly enrolment and schooling is not necessarily translating into learning (Banerjee et al., 2016). In fact, there is a nation-wide call for "catch-up" interventions (ASER Centre, 2022) to address the staggeringly low reading levels.

1.1.1 Language of Learning – a 'Double-Learning Disadvantage'

Many children in India directly join Grade 1 as it is not mandatory to complete pre-primary education before enrolment into a primary school. This means that until school admission, many children from low-income backgrounds are not exposed to English language instruction. However, since they begin their education in English medium schools, children are forced to learn to read in two languages at least – English, which is not their first language, and Hindi or their mother tongue or a state language. Children are expected to read in a language they may not speak or understand, thus causing the "double learning disadvantage" (NCERT, 2021b). In other words, children do not have the reading skills in English language, and they cannot use the language to learn other subjects. This disadvantage continues through their schooling years as children remain under- or un-equipped to read subject textbooks written in English language (Sinha, 2019). Due to this, children are often faced with difficulties with reading and writing and then they are labelled as 'slow learners' (Nag, 2013). The difficulty in learning could be true in case of children's L1 as well because children's both L1 and L2 are likely to be rudimentary at primary level. Also, the opportunities for building awareness of usage patterns of grammar and vocabulary are not adequate even at homes as many parents may not be literate in their L1 (Dixon et al., 2011).

1.1.2 Effective and Adequate Literacy Teaching: A Far Cry

The difficulty in learning becomes more severe with inadequate language teaching practices. In a study coordinated by the National Council of Education, Research and Training (NCERT), India, to assess the status of teaching English at primary level (Dutta & Bala, 2012), it was found that many teachers resort to textbooks and focus less on listening and speaking skills. This holds true even now as children spend large amount of time writing down words or copying answers from the board without knowing what they are copying, and no feedback is provided on the meaning of words upon completion of the task (Sinha, 2019). Emphasis is mostly on rote repetition as the expectation of students is to learn words through repeated contact and teacher-led textbook-based lessons (Menon et al., 2019). Learning by repetition is further undermined by misrepresentation of terms 'letter', 'sound', 'spelling' as teachers use them interchangeably and focus on letter-names and recitation. With blackboard teaching and lecturing as prominent instruction techniques, no room is left for differentiated, remedial, and small group instruction (Bhattarcharjea et al., 2011; Sankar & Linden, 2014).

Furthermore, there is also a lack of effort towards providing differentiated instruction as many teachers share the belief that children coming from disadvantaged economic backgrounds will not be able to learn math or English because they do not have adequate environment and an appropriate attitude towards learning (as cited in Muralidharan et al., 2019). Combined with existing learning level of students, teacher beliefs and knowledge, overreliance on textbooks, and rote and repetition, classroom instruction leads to misalignment between students' learning needs, their ability to engage with instruction, and the instruction itself. This misalignment then causes students to fall behind in language acquisition (Banerjee & Duflo, 2012).

1.1.3 Teachers' Knowledge of Literacy Acquisition

Inadequate classroom pedagogy is often triggered by deficiencies in teachers' knowledge of literacy acquisition. Many teachers often lack the ability to effectively detect errors in students' work and provide relevant support. In a study in Madhya Pradesh and Tamil Nadu it was found that language teachers could identify errors in students' responses in assessments only 50% of time (World Bank, 2016). This challenge has been witnessed in other developing regions such as in Sub-Saharan countries. Bold and her colleagues (2017) documenting teachers' pedagogical knowledge and ability to assess students' learning found that only 1 in 10 teachers had the basic knowledge of pedagogy and none had minimum knowledge of assessments.

Several studies have also indicated that several teachers have misunderstanding of foundational literacy skills (See Hudson et al, 2021 for review). They face difficulty with identifying phonemes in a word and differences between consonant blends and diagraphs (Washburn, Joshi, & Binks-Cantrell, 2011). In a review of studies, Hudson and co-authors (2021) found that teachers struggled with providing appropriate examples during instruction and identifying whether a text is difficult to decode or not. Moreover, often teachers are not cognisant of their actual level of understanding which obscures their awareness and intent to increase their knowledge. Such deficits in understanding of foundational literacy cannot be simply overcome by either providing well-designed resources and materials (Cohen et al., 2017) or through teaching experience alone (Pittman et al., 2020).

Lack of domain knowledge thus makes teacher preparation in foundational literacy imperative as their knowledge of a subject they teach improves students' learning (Glewwe et al, 2013). However, professional development sessions are noted to primarily focus on classroom management and pedagogy related approaches; there is a dearth of domain or subject specific trainings (Menon et al., 2019). Thus, teachers themselves continue to be poor English role models and lack practices for facilitating L2 reading acquisition (Shenoy et al., 2020). As a result, instruction remains limited to a cycle of choral reading, copying question and answers, practicing handwriting and spelling drills.

1.2 Minding the Gaps with Education Technology Combined with Classroom Instruction: A Step Towards Transfer of Learning

EdTech has been recognised as an offering that can increase access to education, address teacher shortages, support learners' learning level, free up teachers from routine tasks, provide adaptive content, support at-risk language learners, and overcome challenges of high teacher-student ratios (Cheung & Slavin, 2012; Escueta et al., 2020). However, despite spendings on education and implementation of EdTech interventions, there is a lack of alignment between the level of classroom instruction and students' learning levels (Glewwe & Muralidharan, 2016) as most developing countries have largely focused on providing hardware without focus on how technology can be effectively integrated into classroom instruction. Studies show that providing instructional material alone or substituting technology-based interventions such as CALs with regular instruction does not improve learning (see McEwan, 2015 for review; Escueta et al., 2020). Efficacy of such EdTech interventions are limited to low to moderate effect on reading outcomes (McEwan, 2015).

While there is mixed evidence for efficacy of CALs, in developing context they have shown positive effects on students' learning gains (Banerjee et al. 2007) and underlined the need for its integration into educational curriculum and along with active involvement of a teacher (Muralidharan et al., 2019; Pitchford, 2015; Mo et al., 2016). In a randomised control trial (Muralidharan et al., 2019) addressing the heterogenous learning levels of students in a class, a game-based CAL software, Mindspark, was implemented with over 3000 children across grades 6-8 studying in a public school in Rajasthan, India. Students received 45 minutes of personalised learning through the CAL and 45 minutes of instruction as per their academic level from a teaching assistant. The effects were assessed using paper-and-pen assessments containing test items with a range of difficulty level – very easy to grade appropriate to assess the intervention effects. The treatment group had large effects in math on test items below grade level and no impact on items at grade level. In Hindi, treatment effects were seen in both cases whereas in English no significant effects were observed. Authors of the study emphasised that for such CAL software to be effective, teachers have an important role in ensuring time of task, supervision of out-of-CAL tasks or homework, and adherence to CAL-based instruction by encouraging attendance. Moreover, in emphasising the gains below grade level, authors of the study underline the importance of students' ability to apply their in-game learning to a new context, in the real world. In other words, students should be able to transfer their learning outside the game, such as on paper (Bainbridge et al., 2021).

Bainbridge (2021) indicated that for in-game learning to be productive and meaningful outside of the game environment, the learning must go beyond "specific transfer" to near or far mixed transfer. She defines "specific transfer" as application within contexts within which learning has occurred. Near transfer is defined as the ability to apply skills in contexts similar to the game context and far transfer in case of contexts which are not similar. Towards this, she points out that learning supports such as "pertaining" and "coaching" within a game can enhance a player's learning process and thereby improve learning outcomes. Pertaining refers to prior instruction that can support a player in learning within the game, whereas coaching refers to orienting players to the learning content instead of other in-game actions. These supports could be helpful in assisting a player especially when they are exposed to new content. However, how, and when should such supports be included in a game for literacy instruction to early grade second language learners in a developing context is a question that needs examination.

1.2.1 Activating Teachers' role: A Promising Alternative

Importance of teachers' involvement to improve learning outcomes in CAL interventions points at two important considerations: one, whether a CAL is implemented during school hours or after, and whether it is reinforcing learning or teaching new content. This was emphasised in an India-based study (Linden, 2008) where the intervention was designed to reinforce teacher-taught lessons through computer-based worksheets in a non-adaptive, self-paced fashion. It was implemented during and outside school hours. Quite paradoxically, the author indicated that computer-based math program implemented within school hours was less productive versus after-school implementation wherein students' test scores increased by 0.28 SD. Linden points out that students who did not miss regular instruction gained more as they had already understood the content in the classroom. This study provided another dimension of integrating teacher and technology – a 'blended learning' environment in which CALs can reinforce teachers' classroom instruction and provide reinforcement to struggling readers (Muralidharan et al, 2019).

Evidence suggests that struggling learners can be greatly supported when teacher instruction and CAL are integrated. In a study, He and his colleagues (2008) compared learning gains when one group was administered an interactive tool PicTalk and another a set of flashcards. The tool allowed children to select a picture with a stylus and hear the word's pronunciation. Also, it allowed hearing instruction not only in English, but also local language, Marathi. Then, to facilitate instruction, teachers were provided with a manual containing activities and drills that they could use with children in both the groups. In this study as well, the poor-performing children in the teacher only group showed smaller treatment effects relative to the group which was exposed to both the interactive tool and teacher-led instruction. In the context of at-risk learners, integration of CAL and teacher thus seems to be a promising alternative to using teacher only or CAL-as-a-supplement approach (McTigue et al., 2019).

When students receive CAL as a supplement and then go back to whole-class instruction, chances are that there is little opportunity for teachers and students to make most of the connection between content in both the settings. Thus, integration of classroom instruction and CAL can bridge the gap between in-game and in-class instruction by for instance, using a shared curriculum. An example of such an integration can be seen in an experimental study (Mo et al., 2016) evaluating 'computers as tutors' in China, comparisons were done between two interventions – one, a game-based English ICT program (CAL) alone, and two, the ICT programme incorporated with regular English instruction. In the second intervention, the programme was implemented by English teachers who were provided with lesson plans and a training session on the implementation protocols. When a new vocabulary was introduced in the ICT program, the English teacher covered listening and spelling exercises for that vocabulary during classroom instruction. Whereas in the first intervention, teachers' role was only limited to the implementation logistics. Interestingly, the first intervention benefited better-performing students more than students who performed poorly at baseline. Whereas, in the second intervention all the students benefited similarly. The authors also posited that since English teachers made the intervention as a part of their regular teaching, incorporation of the programme and its implementation probably required less marginal effort from them.

The virtue of integrating CAL and teacher instruction has been also underlined in a meta-analysis (Conn, 2017) of 56 studies focusing on educational intervention in Sub-Saharan Africa. The author indicates that most effective interventions for improving students' learning entailed a change in instructional techniques, especially those which employed adaptive instruction (d=0.44). The analysis further showed that the integration, or 'computer-assisted strategies', was even more effective than only teacher-led instruction. Similar findings were observed by Wouters and Oostendorp (2013) in a meta-analysis which showed stronger effects when game-based learning (a CAL) was supplemented with teacher-led lessons, and by Clark and co-authors (2016) who point that students' learning improved when teachers' scaffolded instruction. Taken together, these studies point at a key factor for improving learning in India and other similar developing regions: EdTech interventions when aligned to classroom instruction can provide maximum learning opportunity to children and yield most learning gains.

1.3 Literacy Learning Through Phonics Instruction

Literacy acquisition through phonics depends on conjunction of phonology and orthography (Ehri, 2020). Phonology is a system of speech sounds in a language, for instance English has approximately 44 phonemes that can be combined to form larger sound units such as rimes, syllables, and words (Castles et al., 2018). Orthography, on the other hand, include symbols i.e., letters or graphemes which individually or in combination represent phonemes. Phonics instruction explicitly teaches learners about grapheme-phoneme correspondences (GPCs) (Ehri, 2020), for instance the letter j represents /ʤ/ as in jump or letter x represents /k/+/s/ as in box. Knowledge of GPCs in combination with phonological skills such as blending, segmenting can help with pronunciation on unfamiliar printed word, i.e., to decode (Jones & Reutzel, 2012). Phoneme manipulation skills such as blending and segmenting follow a developmental sequence (Holopainen et al., 2020) which can inform phonics instruction as well.

Identifying sounds, for instance first sound in “mug”, is easier than blending individual sound units – /d/ /ɔ/ /g/. Further, segmenting is more difficult than blending, for instance – what are the individual sounds in boat?; and deletion can be even more challenging – what would ‘boat’ be without /b/?

Phonics instruction has a greater immediate and long-term benefit to reading skills than only letter knowledge acquisition and word learning (Ehri, 2020). It can be supportive in teaching English language, which has a complex orthography i.e., connections between graphemes and phonemes are not consistent (Ziegler & Goswami, 2005), for instance word ‘bake’ has /b/ +/eɪ/+/k/; there is no /e/ even though the word ends with e. This irregularity underlines the importance of phonics instruction and emphasises that phonics can help with learning to read and spell unfamiliar and irregular words (Holopainen et al., 2020). Evidence suggests that when letter names and sounds are taught simultaneously children benefit more from the instruction (Piasta et al., 2010). Early readers learn to notice connection between print and speech as and when they begin learning alphabets (Piasta & Hudson, 2022). Letter names thus becomes a foundation for learning GPCs, for which explicit instruction on the association between letter name and letter sound is provided, for instance, letter j is called as “J” and is pronounced as /ʤ/. Developing an understanding of phonics through such instruction can eventually give way to patterns in reading and spelling words (Ehri, 2020). For instance, phoneme /eɪ/, representing long vowel a, has same sound for different spellings such as ai (maid), ay (stay), and eigh (neigh). Equipped with phonological awareness, children can read even unfamiliar words (Ehri, 2020). Thus, the goal of phonics instruction is building an understanding of patterns and generalisations of grapheme-phoneme connections, and not memorising the connections, spellings and words.

Phonics instruction is said to be effective when it is explicit and systematic (Castles et al., 2018). Explicit instruction entails direct and clear instruction and explanation of a concept at hand, modelling of application of the concept, and guided practice combined with feedback (Piasta & Hudson, 2022). On the other hand, systematic denotes an instruction that has a specific scope and sequence, building on prior knowledge and moving from simple to complex skills (Duke & Mesmer, 2019). Further, phonics instruction has two approaches – analytic and synthetic (Castles et al., 2018). Analytic instruction involves teaching phonics using whole words by learning to break them into word parts (Castles et al., 2018). For instance, what’s the common rime between *bit*, *lit*, *chit*? Using this, children might derive the common phoneme /ɪ/ or the common rime /ɪt/.

On the other hand, synthetic instruction involves teaching phonics using parts i.e., GPCs and learning to blend them together to form words (Castles, 2018; Ehri, 2020). For instance, learning to blend phonemes /b/, /ɪ/, /t/ and use the knowledge further to decode words such as lit and chit. Despite differences, both analytic and synthetic methods are said to be effective when implemented explicitly and systematically (Castles, 2018).

Synthetic phonics instruction has been reported to be beneficial for bi- and multi-lingual learners. In a study in 20 low-income private schools situated in Hyderabad, India, Dixon and her colleagues (2011) utilised Jolly Phonics to provide instruction using a synthetic approach. In addition, they used a set of sequential lesson plans which they developed in alignment with the Jolly Phonics material. In this six-month long intervention, the teacher provided synthetic phonics class every weekday for one hour to the experiment group (Grade 1), whereas the control group received as usual instruction, i.e., using rote learning and whole word recognition methods. The results indicated significant improvement in the reading and spelling skills of the experimental group. Classroom instruction showed strong effect (d=1.20) on students' ability to blend three letter sounds and pronounce words. Such findings provide more ground to the idea of utilising phonics instruction in the context of India.

1.4 GraphoLearn: A Computer-Assisted Phonics Instruction

GraphoLearn (GL), originally named as GraphoGame, is a computer-assisted tool that provides explicit and systematic instruction on connections between graphemes and phonemes i.e., written and spoken units of a language. Based on the findings of a longitudinal study – Jyvaskyla Longitudinal Study of Dyslexia (JLD), the tool was originally designed to support children struggling with basic decoding skills in Finnish, a language with transparent orthography (Richardson & Lyytinen, 2014). The JLD study indicated that letter-sound knowledge is a strong predictor of later reading skills. Thus, the game was designed to provide a computer-based learning environment where children can learn to connect graphemes and phonemes. Adapted in multiple languages now, different game versions have been studied in over 20 countries across developed and developing parts of the world to provide and support understanding of correspondences between written and spoken forms of language. In all its versions, players are directed to connections between GPCs in case of transparent orthographies and rimes, blends, and words in case opaque orthographies such as English which has one-to-many GPCs (McTigue et al., 2019). Moreover, in-game content is presented as a spectrum starting with small units, phonemes, and then moving to larger units such as rimes. Though the sequence and quantum of content varies across GL versions.

The scope and sequence of the game is informed by four main theoretical perspectives (McTigue et al., 2019). First, Simple View of Reading (Gough & Tunmer, 1986), presents reading comprehension as a result of decoding skills and linguistic comprehension, thereby emphasising GL's effectiveness in building the relevant skills. Second, psycholinguistic grain size theory which emphasises that availability of GPCs in an orthography vary depending on how transparent or opaque the orthography. In case of English, since larger sound units – rimes – are more consistent than phonemes, exposure to rimes should facilitate reading acquisition (Ziegler & Goswami, 2005). This theory provides basis for exposure to sound units across the spectrum – phonemes, rimes, syllables, and words. Third, Katz and Frost's (1992) orthographic depth hypothesis which indicates that learning to decode in opaque orthographies would take more time because they are more complex than transparent orthographies such as Finnish. Thereby, we see differences in content and its presentation across GL versions. Lastly, the fourth theoretical basis is Ehri's word reading development which happens in phases starting from primary reliance on visual cues and then using letter-sound knowledge in familiar words. This is followed by a completed formation of GPCs and acquisition of decoding skills which later develops a consolidated knowledge of spelling patterns (Ehri, 2005). This informs the in-game instruction focusing on sub-lexical unit (Richardson & Lyytinen, 2014).

In the current study, GraphoLearn English Rime was utilised to provide a computer-assisted phonics instruction as focus on rime is expected to be more beneficial for English learners (Richardson & Lyytinen, 2014). GraphoLearn employs both synthetic and analytical phonics approaches in case of English orthography (Richardson & Lyytinen, 2014). In this mixed-method approach, most consistent and common individual and combination of sounds are presented. First, the game uses visual and auditory stimuli to introduce 6-7 GPCs through repeated exposure. These sub-units are then combined into rhyming word families and finally into words. Soon after, the game provides opportunity to practice the introduced rime units thereby, reinforcing players' recognition of sound units. As players engage with the game, they are also provided explicit feedback either as corrected answer for incorrect responses, or rewards for correct responses. Players' learning level, gauged by their responses, informs progression within the game, thus allowing them to move at their own pace and providing adaptive instruction.

A meta-analysis of 19 studies on GraphoLearn examined the efficacy of the game in several contexts across the world (McTigue et al., 2019). In India specifically, Patel and her colleagues (2018) conducted a randomised control intervention for eight weeks to examine the development in foundational English literacy skills in children who play GraphoLearn. The results indicated that students who played made greater and faster development in in-game letter sound knowledge as compared to students to were not exposed to the game. However, no differences in the learning gains were observed between the two groups of students in oral-and-paper based measures.

In another study (Patel et al., 2021), the authors examined efficacy of GraphoLearn in addition to the relation between in-game progress and students' literacy level pre-and-post intervention. This time, children exposed to GraphoLearn showed greater and faster development in all three in-game measures – letter sound knowledge, rime units and word recognition. However, once again, the children did not show gains on oral-and-paper-based measures. Also, the results indicated that students with more in-game progress not only showed more gains at post-test but also had higher pre-test scores. Yet, students with stronger pre-existing literacy skills in English also did not show effects on measures outside the game. The findings of both the studies by Patel and her colleagues are also in line with other GraphoLearn studies (see McTigue et al., 2019 for review) which show development of sub-lexical skills such as syllable awareness and letter sounds. However, the studies indicate limited to no transfer to lexical skills such as word reading (see McTigue et al., 2019 for review), along with absence of transfer outside the game.

1.4.1 Using GraphoLearn in Classroom Instruction on Phonics

To improve the efficacy of GraphoLearn, studies have emphasized the potential of using GraphoLearn in classroom instruction and increasing involvement of teachers or their interaction with students (Ecohard, 2015; McTigue et al., 2019; Patel et al., 2021). This potential has found substance in the meta-analysis (McTigue et al., 2019) of GraphoLearn studies wherein authors have indicated that high adult interaction had large effect (g=0.47) on students' learning outcomes. Also, Patel and her colleagues (2018, 2022) indicated that lack of GraphoLearn effects were probably related to lack of teacher involvement. It has been suggested that including GL into classroom instruction could aid in greater transfer by providing scaffolding of phonics and giving children opportunity to apply their in-game learning outside the game. This potential of integrating GraphoLearn and classroom instruction for improving lexical skills finds more strength from the finding that high interactions with adults are related to better word reading skills (McTigue et al., 2019).

GraphoLearn has been used as a remedial instruction alongside regular classroom instruction on phonics, in small group or one-on-one settings (Saine et al., 2011). In the study, Saine et al. (2011) provided GraphoLearn as a computer-assisted remedial instruction to struggling readers of Finnish. The children, in groups of five, received teacher-based reading instruction using phonics combined with 15-min individual GraphoLearn time at the beginning of the instruction. This computer-assisted learning remedial group was compared to another group which received only phonics-based reading instruction from their teachers. Here, exposure to the game was included as one of the activities in the reading instruction to remedial students. Results indicated that students who received GraphoLearn and the reading instruction made significant gains. After the intervention, at-risk students were at par with those at grade-level in their letter knowledge, fluency skills and performed better than the reading instruction only group. Also, the at-risk students outperformed grade-level students in reading accuracy. This study emphasises thus that at-risk students would benefit more from classroom instruction combined with computer-assisted learning and need more time to reach at grade-level.

GraphoLearn has also been used to support teachers' literacy skills so that literacy instruction can be correct, appropriate, and adequate (Folotiya et al., 2014). The authors examined the effectiveness of GraphoLearn for developing literacy skills in ciNynaja, one of the official languages of Zambia. One of the three intervention groups had both teachers and students play the game and another two groups had only teachers and only students play the game. In these groups, it was expected that the in-game instruction on ciNynaja literacy would directly or indirectly impact students' learning level. GraphoLearn here worked as a learning tool to directly train students and as a training tool for improving teacher knowledge which would indirectly impact classroom learning level. Results indicated significant improvements in the group where both teacher and student played GL. Students' letter sound knowledge improved and predicted their decoding skills. Improvement in students' language skills was also observed in another study (Ngorosho, 2018) in Tanzania. Most learning gains were observed when both teachers and students were exposed to GraphoLearn.

In the above studies, a triangle with students, computer-assisted learning, and classroom instruction from teachers at its three points indicate a potential template for using technology and classroom instruction to improve student learning levels (McTigue et al., 2019). The current study builds on this template by combining exposure to GraphoLearn with classroom instruction on phonics.

2. THE PRESENT STUDY

The current study, a first of its kind, aims to examine whether GL English Rime, when provided along with classroom instruction on phonics, can improve foundational literacy skills of bi- or multi-lingual students in an English-medium public school in India. Students were exposed to the intervention wherein they played GraphoLearn and studied English phonics from their class teachers. In this intervention, unlike previous studies, the content in GraphoLearn was integrated into the classroom instruction to prevent students from the playing the game in isolation and to facilitate transfer of in-game learning to oral and paper-based measures. Aiming for an effective phonics instruction, the intervention provided systematic and explicit instruction to students (Piasta & Hudson, 2022). Following the content of GraphoLearn allowed for systematic instruction and scaffolding of sound units along with direct explanation and feedback from teachers on phonological awareness skills such as blending supported with explicit instruction. This instruction was enabled by providing teachers with a set of lesson plans which included instructional details along with content. The teachers were not equipped with English phonics knowledge and the necessary instructional skills for teaching phonics. The intervention thus examines the efficacy of GraphoLearn combined with classroom instruction on phonics in an educational context in India wherein students are struggling readers and teachers do not possess pedagogical skills and knowledge of phonics.

The study is built on prior research and recommendations on including teachers and classroom instruction for a comprehensive instruction on literacy skills (Cheung & Slavin, 2012; see McTigue et al., 2019 for review; Patel et al., 2018; 2021; Saine et al., 2011). The importance of including teachers is underlined in the metaanalyis (McTigue et el., 2019) which showed strongest effects (g=0.48) for interventions with adult interactions. Based on the above-mentioned findings, the current intervention is expected to result in the in-game learning gains, especially in sub-lexical skills such as letter sounds and sound units. Since prior studies (Patel et al., 2018, 2021) have raised questions on the transfer of skills from within to outside the game without teacher instruction, the current study might provide answers to the questions with oral-and-paper based measures. The research questions of this study were –

1. How do children who receive GraphoLearn along with GraphoLearn aligned classroom instruction on phonics from class teachers perform compared to students who only receive the GraphoLearn aligned classroom instruction?
2. How does students' attendance to the GraphoLearn aligned classroom instruction influence their foundational literacy skills?

3. METHODS

In this randomised control trial study, the experimental group was provided GraphoLearn combined with classroom instruction on phonics which was aligned with the content in the GraphoLearn. The control group was only provided the classroom instruction on phonics and a math game to control for exposure to technology. The groups comprised of Grade 2 students studying in an English-medium government school in India. This study included pre-assessment using measures on phonological awareness and reading skills for baseline. The baseline testing was followed by the intervention. Then, to test learning gains through this intervention, post-assessment was done using the same set of measures. The data collection and intervention started on 1 July 2022 and ended on 9 September 2022. Permission to conduct the study was taken from Directorate of Education and principal of the school.

3.1 Participants

Out of 80 Grade 2 students, parents of 63 students provided consent to participate. However, 3 students dropped out at the beginning of the intervention due to irregular attendance. Ultimately, 60 students along with their two class teachers participated in the study. Selected students were then divided into two groups of 30 – experimental group and control group. The experimental group received both the GraphoLearn game and classroom instruction on phonics, and the control group received a math game (Math Kids) and the classroom instruction. While forming these groups, students were matched for age and gender. At the end of the study, four students from the control group and 2 children from the GraphoLearn group were removed during analysis because their in-game assessment post-test data was missing from the server. The final analytical sample consisted of a total of 54 students – n (experimental group) = 28; n (control group) = 26.

All students came from low-income families with Hindi as the most common language of communication. Three students were also exposed to regional languages Maithali, Bengali, and Garhwali respectively (See Table 1). All students were learning English as their second or third language in the school and attended after-school private tutoring for English. However, not every student in the class had the same amount of exposure to English language in the school as children were provided admission based on their age. As per the Right to Education Act 2009 (RTE) in India, public school authorities cannot deny a child above 6 years of age admission into the school if the child has not completed his or her elementary education. Thus, some children had not received education at all before their admission in Grade 2.

The classroom instruction on phonics was provided by two teachers. They were trained in instruction for primary classes and conducted lessons for all the subjects in the grade. Like students, they were non-native speakers of English. Their participation in this study entailed attending workshop on foundational literacy skills, playing GraphoLearn, and conducting GraphoLearn-aligned classroom instruction.

3.2 Procedure

The intervention was implemented for 40 days over a period of 7 weeks (July to September) during regular school hours (8.00-13.00). Both the experimental and control group received the intervention at the same time to ensure they received it under same conditions. Before the start of the intervention teachers were provided two, 2-hour workshops on foundational literacy skills during the school hours. In the workshops, they were introduced to the relevance of phonological awareness, grapheme-phoneme correspondences (GPCs), blending phonemes to form rime units and words, and segmenting rimes units into phonemes.

After the start of the intervention, during the game time, both the groups were brought into a separate classroom where they sat next to a device. The researcher opened students' profile for them to ensure that they do not select any other profile by mistake. Each student played under his or her profile name on either GraphoLearn or Math Kids for a duration of 20 minutes on a mobile device equipped with headphones. During each game session, the researcher was present to address technical issues and ensure intervention fidelity. Also, before every session, the researcher instructed students to listen to the audio inputs carefully before clicking any on-screen text, and at the end of the session, students were asked log out of their profile. The researcher went about the room to check whether students needed any support and ensured that they remained on task. The mean playtime recorded at the end of the study was in the range of 10.26–11.18 hours (see Table 1).

Table 1 *Group characteristics*

Characteristic	Experimental (n=28)	Control (n=26)	
	n	n	
Gender			
Male	14	13	
Female	14	13	-
Home Language			
Hindi	27	24	
Hindi+Maithli	0	1	
Hindi+Bengali	1	0	
Hindi+Garhwali	0	1	
Exposure to tuition for English			
Yes	21	22	
No	7	4	
Age (Years)	6.75	6.81	$t(52)=.45$, $p=.65$
Play time (hrs)	11.178(1.37)	10.26(2.24)	$t(52)=-1.843$, $p=.07$
GL-aligned classroom instruction time (days)	20.79(2.53)	18.96(4.26)	$t(40.09)=-1.9$, $p=.06$

All students were exposed to classroom instruction on phonics three to four times a week for 30 minutes in their own classroom during their English language period. The content of the classroom instruction was aligned to the content in GraphoLearn. Lesson plans including the content were provided to the teachers who then delivered the lessons to the students. Classroom instruction on phonics was also attended by the researcher to ensure intervention fidelity. Over a period of 7 weeks, 26 sessions of classroom instruction on phonics were delivered. In this study, classroom instruction on phonics has been reported as the students' attendance to the sessions. At the end of the intervention period, students had attended an average of 19.91 instructional sessions. In other words, students attended the sessions for an average period of ~20 days. There was no significant difference in the number of instructional sessions attended between the experimental and control group (see Table 1).

During the intervention teachers were provided continuous mentoring and support on classroom instruction on phonics. After every session, feedback on their pronunciation of sound units was provided to them. During this feedback, they were notified about the sounds they pronounced incorrectly, and correct pronunciation was shared.

3.3 GraphoLearn

GraphoLearn (GL), formerly known as GraphoGame, is a research-based, computer-assisted tool that trains players on the connections between spoken and written language. In the game, players create an avatar after which they are presented with a letter or letter string and its corresponding sound unit and. Players are expected to correctly match the target sound with its corresponding written form. Players receive the visual input on their individual tablet screens and auditory input through their plugged-in headphone. GL English Rime consists of 25 play streams and 7 assessment streams in which students are provided adaptive practice on letter sounds, rime units and word recognition through various practice levels and assessments. In the game, players are first introduced to single grapheme-phoneme correspondences which are then blended into rimes. Next, the teaching sequence show blending of onsets into rimes to create words. The order in which letter sounds, rime units and words are introduced progress from most orthographically consistent and frequent units to less consistent and frequent units. This order is followed in each of the streams in the game. This sequence of teaching is based on the role of orthographic rime units in English reading development (Kyle & Richardson, 2013). The game provides practice in rhymes across 11 streams and word formation across 15 streams, supporting phonological awareness and spelling skills.

The game begins with assessment 1 after which players, based on their existing skills, are taken to a series of streams and levels which give explicit instruction and practice on individual and a combination of grapheme-phoneme connections. These connections are then used to form rime units and words. Players are expected to identify the correct letter-sound correspondence out of 2 to 7 options. A correct choice takes them to next set of options and incorrect choice provides them with immediate feedback after which they can correct their response. Players should be able to score a minimum of 80% on each level within a stream to move to the next stream. If they are unable to score 80% and above, then players are provided more individualised training on the target sounds they did not answer correctly. They are provided with stars and coins as rewards to boost their motivation. Game data is automatically saved to an external server provided that the device has an active internet connection.

3.4 GraphoLearn-aligned Classroom Instruction on Phonics

All students received GraphoLearn-aligned classroom instruction for an average of 19.91 days (see Table 1). The measure is used to report students' attendance to the instruction sessions. GraphoLearn-aligned classroom instruction was delivered according to a set of lesson plans developed based on the content provided in the GraphoLearn game. These lesson plans incorporated content provided across 25 streams in the game while following the sequence of the content as well. In all, 26 lesson plans were developed and provided to the teachers who delivered one lesson plan per day. Each lesson consisted of 6-10 sounds and/or rime units, including both new and previously learned sounds. These sound units included letter sounds, blends, and rime units systematically sequenced based on difficulty level, frequency, and consistency. In a lesson plan, some sound units were repeated to facilitate revision of sounds. In the lessons, the sequence of sound units provided in the GraphoLearn was followed to ensure that content used is research-based and widely tested.

To facilitate GraphoLearn-aligned classroom instruction, the lesson plans included both a synthetic and analytic approach which then teachers used in the instruction. The synthetic and analytic approach was embedded in the activities provided in the lesson plans. The activities entailed instruction on phoneme and rime identification, blending of phonemes into blends and rimes, and then blending rimes into words (mostly CVC words). Since the content in GraphoLearn is provided in this sequence, the lesson plans also included the sequence and teachers were asked to follow the same as well. Examples of words with target sounds were included in the lesson plans to contextualise the letter sounds, for instance, for letter i, examples such as fish and kite were provided to clearly explain the target sound in the lesson plan. Such examples were provided because over classroom observations it was noted that the teachers were unfamiliar with long and short vowels and got confused while providing instruction on letters with multiple sounds such as /I/ in fish and /aI/ in kite. Also, it was observed that students got confused when the teachers asked them for examples of words with the target sounds but teachers did not know how to address this confusion. Thus, lesson plans were developed in an ongoing manner while accounting for such confusions during the classroom instruction on phonics (see Appendix 4 for a sample lesson plan.).

The classroom instruction on phonics was implemented by one lead teacher and a co-teacher in the Hindi language as the students did not understand English language. The lead teacher instructed the whole class and the co-teacher assisted with classroom management. Both teachers adapted the given lesson plans to students' existing learning level and class environment. The instruction also included the use of teaching aids such as audio clips of target sounds, flash cards, and picture books. Audio clips of target sounds were recorded by the researcher and shared with teachers before the instruction to rule out inconsistencies in the utterance of letter sounds during the instruction. To prevent any errors and gain familiarity with sound units, teachers were asked to deliver the instruction only after practicing the target sounds on GL and after listening to the audio clips. Both teachers played GraphoLearn to completion prior to the end of the intervention.

During the GraphoLearn-aligned classroom instruction, teachers explicitly demonstrated letter sounds by emphasising lip and tongue movement, and by presenting the graphemes on the blackboard. Students were then asked to repeat the sound following the demonstration. Teachers also played audio clips of phonemes as they taught a new GPC or revised one. Instruction on target letter sounds was followed by sounding out letters to form rime units and words. For instance, /*n*/ phoneme was articulated as /*nnn…*/, and /æ/ and /*p*/ were pronounced as /*æææ …*/ and /*ppp…*/ respectively. Then, these sounds were combined to form /*nnn æææ ppp*/ or /nap/. This blending exercise was done successively. To further support instruction on GPCs, CVC words, which were familiar to students and contained target sounds, were used to show how sounds come together to form a word. Towards this, words were sounded out repeatedly to identify a target rime or phoneme. Both for the demonstration and playing audio clips, teacher used a mic, making sure the letter sounds are audible to all the students. Also, audio clip of the target sound was played for students to ensure correct instruction of sounds.

3.5 Math Game

The control group played game Math Kids for an equivalent amount of time as the GL group (see Table 1). Using this app in the intervention ensured that both experiment and control group spent equivalent amount of time using a technology outside their classroom. The math game provides practice in basic mathematical operations – addition and subtraction. Content for both operations is divided into 3 parts arranged in the increasing order of difficulty. In case of addition, first part focuses on counting of numbers, second presents puzzles on addition and last part focuses on addition of numbers. Similarly, in subtraction, first part introduces comparison of numbers, second on puzzles on subtraction and lastly, subtraction using numbers. After playing the 3 parts, students are presented with a quiz and then additional practice on addition and subtraction. Students can play at 3 different levels which allow operations with numbers from 1-5, 1-10 and 1-20. For grade 2 students, the level with numbers 1-20 was selected. Players are not allowed to move forward in the game unless they select a correct response out of three options. Like GL, the game provides visual stimuli in case of an incorrect response and rewards players with a sticker or stamp or a gift after every 4-5 correct responses. Also, players play under their own profile and their performance is stored in the app. To ensure that the control group did not receive any visual or auditory input in English, the in-game content language was set to Hindi.

3.6 Measures

Efficacy of the intervention was determined by assessing students at pre- and post-test level using in-game assessments in GL and experimental oral-and-paper based assessments. The GraphoLearn in-game assessment constituted 3 tasks: letter-sound knowledge, rime unit recognition, and whole word recognition. Out-of-game measures, adapted from DIBELS – Dynamic Indicators of Basic Early Literacy Skills 8th Edition, (University of Oregon, 2021) and ASER – Annual Status of Education Report, India (ASER Centre, 2014), were used to check transfer of skills from GL and development of foundational literacy skills through teacher instruction. Students completed 7 paper-based tasks – letter names, letter sounds, DIBELS word reading, GL word reading, words similar to GL word reading, non-word reading, and oral reading fluency; and 4 oral tasks – initial phoneme identification, last phoneme identification, phoneme blending, and phoneme segmenting. In addition, students completed a paper-based vocabulary measure – PPVT5 (Peabody Picture Vocabulary Test 5th Edition) – at pre-test only.

3.6.1 In-game measures

All students in the study played the three in-game assessments at pre- and post-test. In the GL letter-sound task, children play 24 trials of letter sounds. Each trial requires children to pick the correct letter, out of six to seven options, corresponding to the sound presented to them. In GL rime unit recognition, children play 24 trials where they match target sound unit with the correct two-to-four letter string out of six options displayed on screen. Finally, in the word recognition task, children play 47 trials wherein in each trial children match the correct word, from a set of five on-screen options, with the sound unit provided as an auditory input. Unlike the letter sounds where children play 24 trials irrespective of number of correct responses, the rime unit and word-recognition tasks discontinue if more than 50% of responses are incorrect.

3.6.2 Oral-and-Paper-Based Measures

All students in the study completed oral-and-paper-based assessments. As mentioned, all assessments were experimental except for one norm-referenced measure (PPVT-5) and one criterion-referenced measure (Oral Reading Fluency). For all assessments, the score was calculated as the number of correct responses. Experimental measures were used to ensure the relevance of the test items to the context of the children. Use of standardised measures could have limited assessment of intervention-specific knowledge domains and development of literacy skills given the learning level of children (Jim et al., 2021). Also, measures such as letter sound knowledge and GraphoLearn word reading were used to assess the transfer of skills from in-game to out-side-the game for a set of trained items. In a similar vein, the measure assessing reading of words similar to word items in GraphoLearn was utilised to assess whether students can generalise their literacy skills from trained to untrained items.

Vocabulary Measure

PPVT-5 is a norm-referenced measure for English language that measures receptive vocabulary knowledge (Dunn, 2019). It comprises of two forms – Form A and Form B, each consisting of 240 items ordered by their difficulty level. Each PPVT item consists of a stimulus word and an assortment of four pictures, one depicting the word and remaining three as distractors; all words are balanced for colour, size, and orientation. (Dunn, 2019). The task requires student to give a one-word oral or a non-verbal response. Words are taken from different parts of the speech – noun, verbs, adjectives – and content categories. While different start points based on age are provided in the assessment booklet, the test was administered from item 1 up till item 76, which is the last item for age group 7 years 0 months to 7 years 11 months, the maximum age of the sample. For each correct item, the student received a score of 1 and for an incorrect item, a score of 0 was given. Conducted as per the guidelines provided in the PPVT5 manual booklet, students were assessed from item 1 until they reached the ceiling that is responded incorrectly to 6 consecutive items. In case, 6 consecutive score of 0 is not achieved, the last item (item 76) is considered ceiling. Total number of correct responses out of 76 was the outcome score.

Letter Name and Letter Sound Knowledge

Adapted from DIBELS 8th Edition Letter Naming Fluency task for Grade 1, Letter name knowledge measure and Letter sound knowledge measure consisted of 54 randomly ordered letters each, both upper and lower case. (See appendix 5 for the measure.) The tasks were administered one-on-one. Students were given an A-4 sized paper and instructed to read the letter names aloud in case of letter name knowledge and share the sound for each letter in case of letter sound knowledge. The test administrator pointed to letter using index finger to direct students to the next letter on the task sheet after their response. Each correct response was given a score of 1 and an incorrect or no response was given a score of 0. Cronbach alpha for the letter name knowledge measure was .98 at pre-test and .98 at post-test. Cronbach alpha for the letter sound knowledge measure was .84 at pre-test and .97 at post-test.

Phoneme Identification

Students were tested on initial phoneme identification (IPhI) and last phoneme identification (LPhI) using two oral measures consisting of 10 words with 3-5 letters. Words used in the measures were taken from DIBELS battery for Grade 1 students (See appendix 6 for IPhI and LPhI measure) as it aligned with the vocabulary introduced in Grade 1 and Grade 2 level English textbook used in the school. In case of IPhI measure, the child was instructed as follows: "I am going to say a word. After I say it, you have to tell the first sound that you hear in the word. Now, tell me the first sound in *cot*. Correct response from a child was followed by – "Good, the first sound in '*cot'* is /*c*/. Incorrect response was followed by – "the first sound in '*cot'* is /*c*/. Now, your turn. What is the first sound in '*cot'*?" After demonstrating a sample item to ensure the child understands the task, the test administrator conducted testing. Similar instruction was given in LPhI task. To ensure each child fully understood the tasks, instructions were also given in Hindi when required. In case of no response from the child for 6 consecutive seconds, the test administrator said the test item again and then moved to the next word. In both tasks, all 10 words were presented to children. Each correct response was given a score of 1 and an incorrect response, 0. Cronbach's alpha for IPhI measure was .24 at pre-test and .85 at post-test. Cronbach's alpha for LPhI measure was .68 at pre-test and .89 at post-test.

Phoneme Blending

Experimental oral measures for phoneme blending (PhB) and segmenting (PhS), each consisting of 10 CVC words, were administered to all the students at both pre- and post-testing. (See appendix 7 for PhB and PhS measure). In case of PhB measure, children were instructed as follows: "I am going to say three sounds one after the other. After I say them, you will combine the sounds to form a word and then say the word aloud. So, if I say /*b*/, /*a*/, /*t*/, you will say '*bat*'. Now, tell me what will /*m*/ /*a*/ /*t*/ become when put together?" Correct response from a child was followed by – "Good, the three sounds /*m*/ /*a*/ /*t*/ form the word '*mat'*." Incorrect response was followed by – "Sounds /*m*/ /*a*/ /*t*/ when added form the word '*mat'*. Now, your turn. What word will we get when we blend sounds /*m*/ /*a*/ /*t*/?" This set of instructions was followed by administration of test items.

In case of no response for 6 seconds, the test administrator said the test item again and then moved to the next word. During the task, students blended sounds orally. The instruction for the task was adapted from the DIBELS battery. To ensure each student fully understood the task, instructions were also given in Hindi when required. In the measure, each correct response was given a score of 1 and an incorrect response, 0. Cronbach's alpha for PhB measure was .84 at pre-test and .86 at post-test.

Phoneme Segmenting

Experimental oral measure for phoneme segmenting (PhS) consisted of 10 CVC words and was administered at pre- and post-test. (See appendix 7 for PhS measure). Children were instructed as follows: "I am going to say a word. After I say it, you have to tell me all the sounds the word contains. For example, if I say '*bat'*, you will say /*b*/, /*a*/, //*t*/. Now, tell the sounds in '*cat'*." Correct response was followed by – "Good, the sounds in '*cat*' are /*c*/, /*a*/, /*t*/." Incorrect response was followed by – "The sounds in '*cat'* are /*c*/, /*a*/, /*t*/. Now, your turn. What are the sounds in the word '*cat*'?" This set of instructions was followed by the administration of test items. In case of no response for 6 seconds, the test administrator said the test item again and then moved to the next word. During the task, students segmented sounds orally. The instruction for the task was adapted from the DIBELS battery. To ensure each student fully understood the task, instructions were also given in Hindi when required. In the measure, each correct response was given a score of 1 and an incorrect response, 0. Cronbach's alpha for PhS measure was .45 at pre-test and .90 at post-test.

Word Reading

Three word-reading and one pseudo-word reading paper-based measures were administered to all the students. The first word reading measure (DWR) consisted of 20 test items taken from DIBELS battery for Grade K since DIBELS Grade 1 battery level was too advanced for children based on their textbooks. The second word reading measure (GWR) comprised of 20 words from GraphoLearn, and third measure (SGWR) comprised of 20 words which are phonologically similar to the words in the GWR task. In both the measures, test items progressed from 3 to 7 letter words which tested transfer of skills from digital game-based to paper-based environment. The fourth measure (NWR) – non-word reading consisting of 20 test items – was adapted from DIBELS grade 1 battery for Nonsense Word Fluency to assess pronunciation of words. (See appendix 8 for word reading tasks.) All four reading measures were administered one-to-one with the child and scored out of a maximum of 20 – 1 for correct response and 0 for incorrect or no response. Children were given an A-4 sheet with printed words and asked to read one word at a time and move forward if they are unable to read it.

Cronbach's alpha for the four reading measures was in the range of .84 - .90 at pre-test and .86 - .90 at post-test. Since the measures were highly correlated with each other, they were combined to form the Composite word reading (CWR) measure. Cronbach's alpha for CWR was .96 both at pre- and post-test level. All four reading measures and the composite word reading measure are used for the analysis towards first research question.

Oral Reading Fluency

ASER's reading task (a story – for testing literacy level in India) was used to assess oral reading fluency (ORF) at pre- and post-test (See appendix 9 for the measure.) The tool is criterion referenced and is included in a 4-part test used to assess children's basic reading skills in India. The test's inter-rater reliability is reported to be .64 (Vagh, 2012). Level of the story corresponds to Grade 2 level within Indian context, and thus, was used as it is for testing. In the task, the child was provided an A-4 sheet with the text and instructed to begin and end reading the paragraph at the start and end of 1-minute time limit. The test administrator kept the time and noted the number of words read correctly along with the time taken in case it was less than one minutes. Here, a total of correctly read words is taken as the raw outcome score.

3.7 Data Analysis

Data collected using above-mentioned measures were collated and analysed using SPSS version 28.0. For the oral-and-paper based measures, the scores were manually entered, and for GL measures, the scores were downloaded via the game server. In addition to the scores, play time and attendance for classroom instruction on phonics were also recorded for both experimental and control group (see Table 1). Both groups were then compared with each other on age, play time, and attendance to the classroom instruction on phonics using *t*-test to check whether there were any significant group differences. No group differences were found.

Prior to further analysis, pre- and post-test scores were checked for normality. All three in-game measures – GL letter sounds, GL rime-units and GL word recognition resembled a normal distribution at both pre- and post-test. In the case of the oral-paper-based measures, extreme outliers were found in case of letter name knowledge, letter sound knowledge, phoneme segmentation, non-word reading, and oral reading fluency at either pre- or post-test. While these outliers were winzoried, only pre-test letter name knowledge and pre-test oral reading fluency approximated normal distribution and the remaining three measures did not. For other oral-and-paper based measures, assumption of normality was met. The oral vocabulary test PPVT-5 only administered at pre-test also approximated normal distribution.

All the variables were then further examined to answer the research questions. For the first research question, both parametric (independent sample *t*-test) and non-parametric (Mann-Whitney *U* test) tests were used to examine group differences at pre- and post-test. This was also done due to the small sample size. However, as the results did not differ from those given by the *t*-test, the *t*-test results are reported (see Table 2 and 3). For the non-normally distributed tasks as well, both the *t*-test and Mann-Whitney U test was used. The results did not differ in case of phoneme segmentation and non-word reading pre-test measures. Thus, only *t*-test results are reported for these tasks as well. Mann-Whitney U test results are reported for only letter sound pre-test measure as the results differed from *t*-test results. Effect sizes accompanying the group differences are reported for the parametric tests using Cohen's *d* criteria: $d \leq .2$=small effect; $d \geq .5$=medium effect; and $d \geq .8$=large effect (Cohens, 1988).

To further examine differences in groups' skill development, repeated measures ANOVA was used to compare the effects of time (development between pre-test and post-test), group (effect of being in the experimental or control group), and group*time interaction (group differences in change over time) on development in students' scores. In addition to the normally distributed measures, repeated measures ANOVA was used for the non-normally distributed tasks as well. This was done because gains in letter-sound knowledge was one of the key outcomes targeted through the intervention. Also, a highly skewed distribution of scores and floor effects were seen on letter sound pre-test and phoneme segmentation pre-test, indicating low to non-existent skill level. However, the scores in both the tasks at post-test showed variation and met the assumption of normality, thereby suggesting gains in letter-sound knowledge and phoneme segmentation skills. Similar gains were observed in case of non-word reading and oral reading fluency pre-test measures whose Mann Whitney U results also did not differ from its *t*-test results. Since the first research question aimed to examine the effectiveness of GraphoLearn combined with GraphoLearn aligned classroom instruction, repeated measures ANOVA supported with examining the development of literacy skills across all the measures.

To examine how students' attendance level of the classroom instruction on phonics predicts gains in their phonological awareness and reading skills, first Pearson's correlations were estimated to find relation between classroom instruction on phonics and in-game and oral-and paper-based post-test scores. Significant medium correlations were identified between the instruction and three oral-and-paper based measures – letter names, letter sounds and phoneme blending. Next, a hierarchical regression analysis was conducted to examine the unique effects of GraphoLearn-aligned classroom instruction on the measures. In the hierarchical regression predictor variables are added in steps, thereby allowing the examination of the unique effect of each variable(s) over and above the previously entered variable(s). In the current study, the variables were entered into the model in three steps to predict letter name knowledge, letter sounds knowledges and phoneme blending.

In all the models, group (control and experimental), vocabulary, and pre-test measure corresponding to the post-test measure were used as control variables. To examine whether GraphoLearn-aligned classroom instruction can predict letter-name post-test, first, group and vocabulary were entered first into the model to control for the variance emerging from differences in the group to which children belonged and their pre-existing vocabulary skills. The letter-name pre-test score was then entered into the model to control for the variance emerging from differences in pre-existing letter-naming skills. Lastly, GraphoLearn-aligned classroom instruction was entered into the model allowing the examination of the unique contribution of the instruction to letter name knowledge over and above the control variables. Similarly, in case of letter-sounds post-test, groups and vocabulary were added first in the model. Then, letter-sound pre-test was added second, and in the third step, GraphoLearn-aligned classroom instruction was added to the model. Same method was followed in case of phoneme blending post-test. This design allowed the examination of the contribution of the instruction to letter sounds knowledge and phoneme blending respectively, while controlling for the effects of the group, vocabulary skills and their pre-test scores.

3.8 Intervention Fidelity

The intervention comprised of exposure to GraphoLearn, a computer-assisted learning tool, and classroom instruction on phonics. The experimental group was exposed to both GL and the instruction and the control group was exposed to a math-game and the instruction. Over a period of two months, this intervention was implemented for 40 days during which the game component was administered every day and the instruction component was delivered two to three times a week. On a game day, three rounds of game sessions were administered. In each round, 20 students, 10 experimental and 10 control, played their respective games. At the end of the three rounds, all 60 students played either GraphoLearn or math game. On the day of the classroom instruction, the instruction was delivered to all the students together in their classroom. In each game and the instruction session, the researcher was present from beginning till the end of the sessions to ensure that the intervention is delivered as per the design.

To keep a track of frequency and duration of the sessions, the researcher took attendance of students during both the game (GL/math) time and GraphoLearn-aligned classroom instruction time. The attendance recorded during all the sessions was manually compiled at the end of the study. In case of the experimental group, GL software also logged the number of days and time spent playing the game for each student. In case of the control group, the game did not record time the spent in playing.

To ensure the quality of delivery of the game component, first 20 mobile devices and headphones were set up in a room. GL was set up on 10 devices for the experimental group and math game was set up on the other 10 devices for the control group. On each device, the player's name was entered by the researcher to ensure students play under their own profile. Students were then called into the room and sat next to the device. After all the students were settled at their places, researcher went about the room to cross check that students' name matched the profile on the device. Then, the students were asked to start the game at the researcher's signal. Both groups were given 20 minutes of game time which was recorded using a timer. At the end of the 20 minutes, students were asked to stop playing the game. The researcher went about the room to help them log out of the game. This procedure was repeated for game sessions throughout the intervention.

To ensure the quality of delivery of the GraphoLearn-aligned classroom instruction, teachers were provided with a set of lesson plans which contained instructions on teaching grapheme-phoneme correspondences, rime units and words that occur in the GL tool. The lesson plans for the GraphoLearn-aligned classroom instruction were developed by the researcher and implemented by the class teachers. The lesson plans followed a standardised format for all of the instruction sessions which began with setting agenda for the day, followed by introducing the letters along with letter sounds in the prescribed order both orally and using the chalkboard, and then asking students the sound units to which they were introduced. Within this format, teachers could choose the method of introducing letter-sounds and the amount of time spent on each sound depending on how their students responded to the instruction.

All of the GraphoLearn-aligned classroom instruction sessions were monitored by the researcher. Each session lasted from 30 to 35 minutes during which the researcher was present. Through structured field notes, the researcher recorded start and end of the instruction time, target sound units covered during the instruction and the accuracy with which they were pronounced, amount of time spent on each sound unit, examples of words where target phonemes and rime units were used, and time spent on classroom management versus on the instruction. Based on the observations, the lesson plans were modified on an on-going basis to help improve instruction.

3.9 Ethical Solutions

Before the implementation, permissions were taken from the state's department of education who had then issued an official notification letter to the school. Then, the permission was taken from the school's principal who then introduced the researcher to the class teachers. Along with the teachers, a parent-teacher meeting was organised, during which students' parents were notified about the study, their rights and their children's rights as study participants. After this, parents were provided consent form, privacy notice and research notification in the local language, Hindi (see appendix 1, 2 and 3 respectively). Over the meeting, it was explicitly mentioned that their participation in the study was completely voluntary and that they could pull out their child without giving a reason. Also, they were told that there'd be no negative consequences for them or their child if they withdraw the child from participation. After agreeing that they had been adequately informed about the study, their rights and privacy of their data, parents returned signed copies of the consent form to the researcher. Since all parents did not come on the day of the meeting, consent forms were taken on the next day as well. After receiving the consent, parents were asked to fill in a demographic survey form to collect information regarding their educational and language background, language used for conversation at home, access to digital media such as mobile phone and television, and access to text material in English at home.

After receiving parents' consent, their children were notified about the study and provided an age-appropriate consent form, which was filled by them and submitted to the researcher. Children were also told that they could say no to participation when they wanted. They were not coerced to participate even after receiving consent from their parents and were not harmed in any way. Children were told that they would play a game and their teacher would teach them English during which the researcher will also be present.

After the intervention the collected data was documented and processed for data security following the JYU guidelines. Data collected through the demographic survey and consent forms was entered and stored on an excel file wherein personal identifier – students' names were pseudonymised using codes. A master-sheet containing students' names and codes are stored in a locked drawer. In addition, the hard copies of the consent form are in a locked drawer as well. They will be scanned and stored in the U-drive by July 2023. The data from demographic survey questionnaires is stored in a .csv and .xlsx file. The hard copies of the questionnaires are also in a locked drawer. Access to the locked drawers is known and available to the researcher. With respect to the quantitative data (students' scores) collected for the study, the data is pseudonymised and stored in .sav and .xlsx format under a username and password protected folder. This data will also be stored on the U-drive to ensure that it is secured against all unwanted contingencies and the data is automatically backed up. Access to the quantitative data is also limited to the researcher.

After the completion of the study, the data will be anonymised to prevent identification of people and merging of new data. The hard copies will be shredded along with the master-sheet file containing names and corresponding codes. All the .xlsx and .csv files containing data which can be combined to identify people will be overwritten to ensure complete anonymisation. In case of indirect identifiers such as age and child's mother tongue, the de-identification of values will be done by replacing the original values with values that are independent of the research participant. To further support with anonymisation of data after thesis publication, generalisation and randomisation techniques will be used. To support with generalisation, numerical values will be modified in scale and categorical values will be recategorized. To reduce accuracy of indirect identifiers, noise will be added to the categorical values. To make sure that the anonymisation is successful, participants' data will be checked to see whether they can be singled out by linking their records or whether their original values can be inferred from the replacement value.

4. RESULTS

The aim of this study was two-fold. One, it aimed to examine the effectiveness of GraphoLearn when it is implemented along with GraphoLearn-aligned classroom instruction. Two, the study also aimed to examine how students' attendance to the GraphoLearn-aligned classroom instruction predicts their foundational literacy skills at post-test. Towards these aims, all students received GraphoLearn-aligned classroom instruction and the experimental group played GraphoLearn and control group played a math game. To meet the first aim, both groups' pre-test scores were examined for any pre-existing group differences in foundational literacy skills using t-tests. After establishing experimental and control group did not differ at pre-test, the groups were compared using repeated measures ANOVA to see their development between pre- and post-test, effects of differences in treatment to both the groups, and group differences in change over time. This analysis allowed us to compare the effectiveness of GraphoLearn combined with GraphoLearn-aligned classroom instruction versus only GraphoLearn-aligned classroom instruction.

Towards the second aim, first as a preliminary examination, relation between all the measures at pre- and post-test was studied using Pearson's correlation. After this, as a first step to identify whether students' attendance to the GL-aligned classroom instruction predicts students' foundational literacy skills at post-test, relation between all the measures (in-game and out-of-game) at post-test and the instruction was examined using Pearson's correlation. Upon finding a significant relationship between the students' attendance to the GraphoLearn-aligned classroom instruction and three oral-and-paper based measures (letter name task, letter sound task, and phoneme blending), hierarchical regression analysis was performed to examine whether the instruction makes any contribution to the development of foundational literacy skills at post-test.

4.1 Pre-test and Post-test Group Comparisons

In line with the first aim mentioned above, an independent samples *t*-test was conducted to examine whether there were group differences at pre-test or post-test. Since the sample size was small, group differences were also analysed using non-parametric measure Mann-Whitney U but as the results did not differ from the *t*-test results for all the measures except for oral-and-paper based letter sound pre-test, only *t*-test results are reported. Whereas for the letter sound knowledge pre-test, only Mann-Whitney U results are reported (see Table 3).

For in-game measures, the results showed that there were no group differences at pre-test and effect sizes were small (*d*=.17-.31) (see Table 2). At post-test, group differences in favour of the experimental group were significant for all the in-game measures; GL letter sounds ($t(52)= 3.70, p<.001$), GL rime units ($t(33.33) = 4.19, p<.001$) and GL word recognition ($t(52)=3.18, p < .01$). Cohen's *d* effect sizes were large as well – GL letter sounds ($d= 1.01$), GL rime units ($d=1.11$) and GL word recognition ($d=.87$), thus emphasising that the differences between experimental and control groups' learning gains were large at post-test.

For oral-and-paper based measures, results showed no significant differences between the groups at neither pre-test nor post-test (see Table 3). At pre-test, Cohen's *d* effect sizes for group differences were small and supported the *t*-test finding of no significant group differences in letter name (.10), last phoneme identification (.16), phoneme blending (.36), DIBELS word reading (.19), GraphoLearn word reading (.10), similar to GraphoLearn word reading (.11), composite word reading (.11) and oral reading fluency (.18). In case of initial phoneme identification (.52) and PPVT vocabulary test (.48), effect sizes were medium. However, group differences were not significant in these measures as well.

At post-test, effect sizes for group differences were again small and supported the *t*-test finding of no significant group differences in letter name knowledge (.10), letter sound knowledge (.25), initial phoneme identification (.12), last phoneme identification (.29), phoneme blending (.31), phoneme segmentation (.08), DIBELS word reading (.08), GL word reading (.05), similar to GL word reading (.09), non-word reading (.04), composite word reading (.11), oral reading fluency (.22). Confidence intervals for all these measures crossed zero at pre- and post-test.

Next, the Mann-Whitney U test results showed significant differences between the groups for letter sound knowledge at pre-test. There were significant differences between control (Md=.00) and experimental group (Md=2.0) for the letter sound knowledge measure (U=254, z=-2.06, p=.04). In case of phoneme segmentation and non-word reading pre-test measures, no significant differences were found between the experimental and control groups and the Mann-Whitney U results did not differ from their *t*-test results. Thus, the Mann-Whitney U value and effect size are reported in Table 3 only for the pre-test letter sound knowledge measure. In case of the post-test letter sound knowledge, the *t*-test and Mann-Whitney results did not differ, and thus only *t*-test results are reported.

In sum, the post-test results for oral-and-paper based measures indicate that experimental group's learning gains did not differ from the control group after the exposure to GraphoLearn and GL-aligned classroom instruction. However, the experimental group performed significantly better than the control group in case of in-game measures. As a next step to examine these group differences further and answer the first research question, both groups' development from pre- to post-test were compared.

4.2 Group Comparisons of Development from Pre-test to Post-test

Repeated measures ANOVA was used to compare the effects of time (change from pre- to post-test), group (GL versus control), and time*group interaction on the scores (group differences in change over time).

In all in-game measures (GL letter sounds, GL rime units, and GL word recognition), significant effects of time, group and group*time interaction were found (see Table 2). This indicates that both groups showed improvement from pre- to post-test (See Figure 1) and experimental group showed significantly higher scores and faster development than the control group in all the measures. In the GL letter sounds measure, there were statistically significant effect of time ($F(1,52)=99.69$, $p<.001$), of group ($F(1,52)=11.60$, $p<.001$), and of group*time interaction ($F(1,52)=12.80$, $p<.001$). In the GL rime units measure, statistically significant effect of time ($F(1,52)=26.13$, $p<.001$), of group ($F(1,52)=94.93$, $p<.001$), and group*time interaction ($F(1,52)=18.65$, $p<.001$) were present. Finally, in the GL word recognition task, there were statistically significant effects of time ($F(1,52)=30.86$, $p<.001$), of group ($F(1,52)=6.55$, $p<.05$), and group*time interaction ($F(1,52)=8.98$, $p<.01$).

In all the oral-and-paper based measures, there was a significant main effect of time (see Table 3), showing improvement in both the groups between pre- to post-test (see Figure 2). However, there were no significant effects of group and group*time interaction on all the measures with both groups showing similar scores in oral-and-paper based tasks and similar pace of development. Despite the lack of group and interaction effects, the gains observed across all oral-and-paper based measures for both the groups were notable because of the context of the intervention. While the intervention group studying in Grade 2 had exposure to English-medium learning for over 12 months, their phonological awareness and reading level was poor, as indicated by the mean scores at pre-test. After the intervention where students in the experimental group received average play time of 10.26-11.18 hours and average exposure of approximately 20 days to the GraphoLearn aligned classroom instruction, their learning gains at least doubled across all the measures. Students in both the experimental and control group showed improvement. This suggests the effectiveness of the intervention and that only GraphoLearn-aligned classroom instruction or when combined with the GraphoLearn tool can lead to learning gains over time. Through this intervention, there might had been transfer of skills from the in-game to outside-the-game measures. Also, transfer from sub-lexical skills such as letter sounds, phoneme blending and phoneme segmenting to lexical skills such as word and non-word reading items could have been possible. Even though students' scores were still low in most of these measures at post-test, the improvement in the short duration of the intervention underlines the efficacy of the current intervention.

Table 2: *Descriptive statistics and group comparisons on GraphoLearn tasks*

Measure	Assessment	Experimental n = 28 *M*(SD)	Control n = 26 *M*(SD)	*t*	*d*	Group effect *F*(1,52)	Time effect *F*(1,52)	Interaction effect *F*(1,52)
Letter sounds	Pre-test	5.25(2.30)	4.54(2.25)	$t(52)=1.14$	.31			
	Post-test	14.86(6.29)	9.08(5.05)	$t(52)=3.70$***	1.01	11.60***	99.69***	12.80***
Rime units	Pre-test	.82(.94)	1.04(1.08)	$t(52)=.79$	.21			
	Post-test	5.39(4.74)	1.42(1.58)	$t(33.33)=4.19$***	1.11	94.93***	26.13***	18.65***
Word recognition	Pre-test	2.04(1.55)	1.77(1.53)	$t(52)=.63$	.17			
	Post-test	4.61(2.69)	2.54(1.98)	$t(52)=3.18$**	.87	6.55*	30.87***	8.98**

*$p \leq .05$, **$p \leq .01$, ***$p \leq .001$

Figure 1 *Group comparisons of development from pre-test to post-test on GraphoLearn tasks*

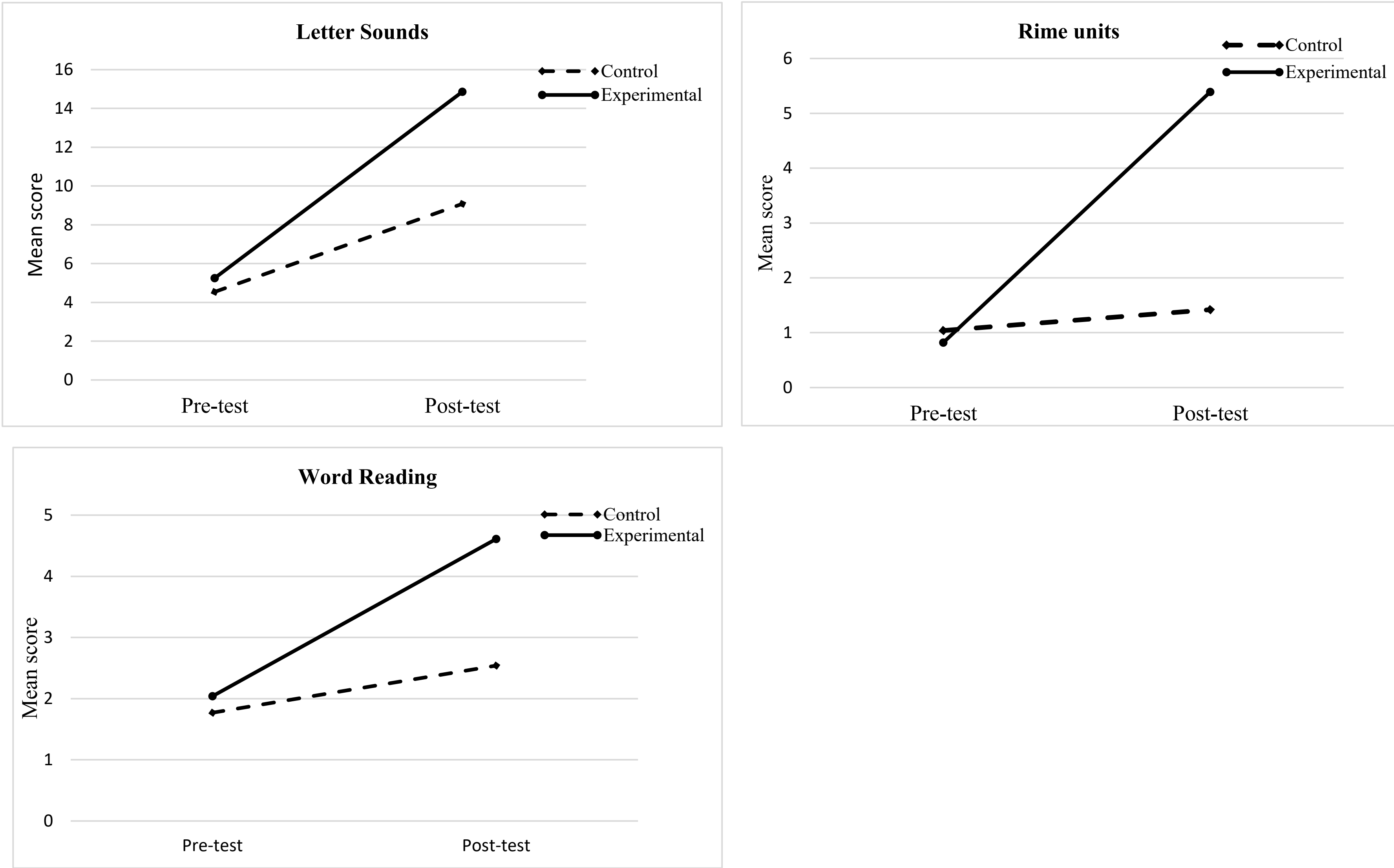


Table 3 *Descriptive statistics and group comparisons on oral-and-paper based tasks*

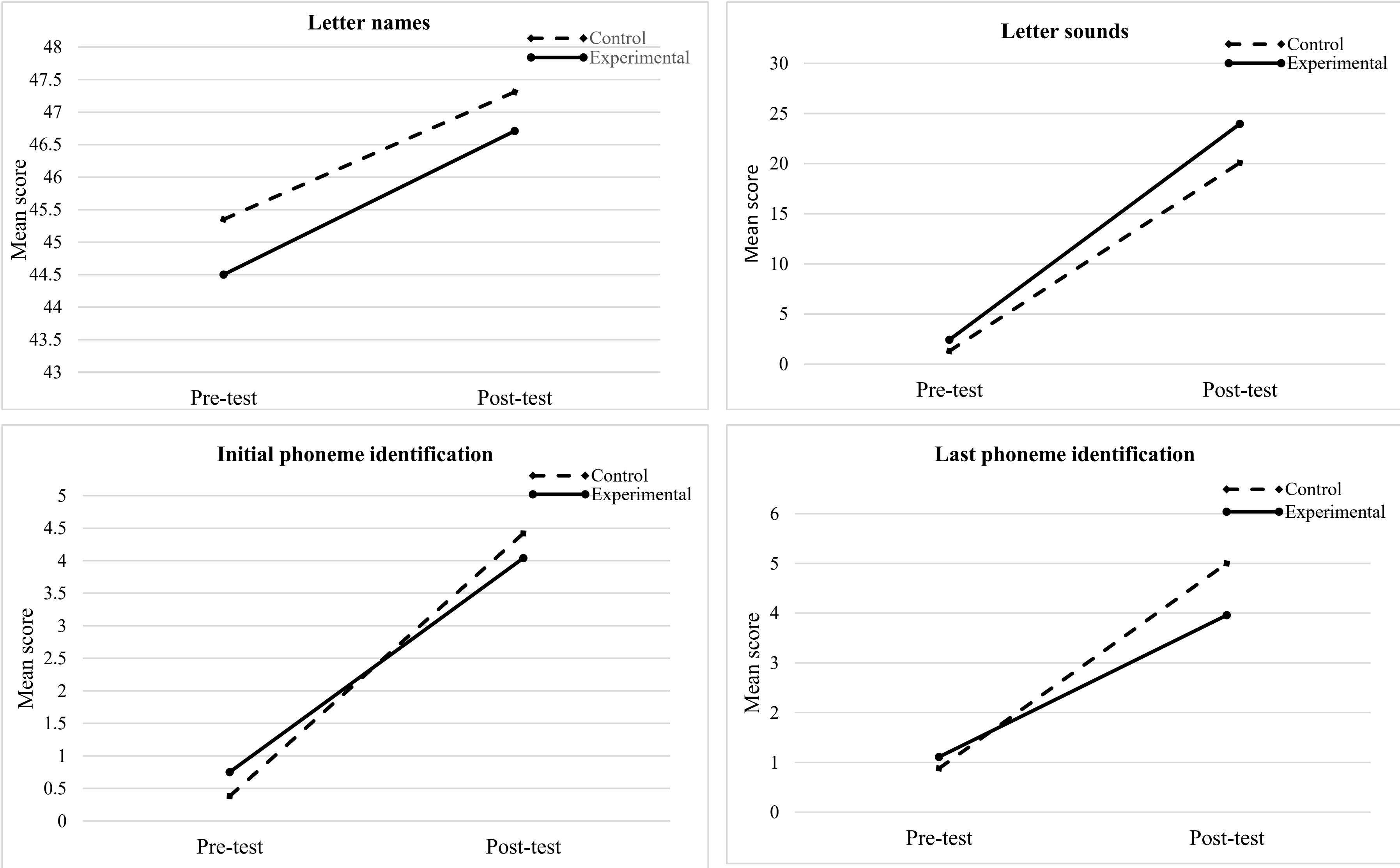
Measure	Assessment	Experimental n = 28 *M*(SD)	Control n = 26 *M*(SD)	*t*	*d*	*U*	Group effect *F*(1,52)	Time effect *F*(1,52)	Interaction effect *F*(1,52)
Letter Name	Pre-test	44.50(8.23)	45.35(8.25)	*t*(52)=.38	.10	-			
	Post-test	46.71(6.09)	47.31(5.91)	*t*(51.89)=.36	.10	-	.14	22.57***	.08
Letter sounds	Pre-test	2.43(2.89)	1.31(3.16)	-	.28[a]	254			
	Post-test	23.96(15.70)	20.08(15.01)	*t*(52)=.9	.25	-	1.25	100.63***	.47
Initial phoneme identification	Pre-test	.75(.79)	.38(.57)	*t*(48.89)=1.94	.52	-			
	Post-test	4.04(3.16)	4.42(3.12)	*t*(52)=.45	.12	-	.00	83.82***	.89
Last phoneme identification	Pre-test	1.11(1.39)	.88(1.45)	*t*(52)=.57	.16	-			
	Post-test	3.96(3.37)	5(3.71)	*t*(52)=1.07	.29	-	.52	55.14***	1.79
Phoneme blending	Pre-test	2.82(2.76)	1.88(2.40)	*t*(52)=1.32	.36	-			
	Post-test	5.86(3.27)	4.85(3.23)	*t*(52)=1.14	.31	-	1.82	74.18***	.01
Phoneme segmentation	Pre-test	.29(71)	.23(.59)	t(52)=.31	.08	-			
	Post-test	3.46(3.39)	3.38(3.37)	*t*(52)=.09	.02	-	.02	51.16***	.00
DIBELS WordReading	Pre-test	2.75(3.30)	2.12(3.40)	*t*(52)=.70	-.19	-			
	Post-test	3.54(4.17)	3.23(3.56)	*t*(52)=.29	-.08	-	.26	7.21**	.22
GraphoLearn WordReading	Pre-test	2.39(3.68)	2.08(2.83)	*t*(52)=.35	-.10	-			
	Post-test	4.50(4.19)	4.27(4.44)	*t*(52)=.20	-.05	-	.08	41.98***	.02
Similar to GraphoLearn Word reading	Pre-test	2.46(3.34)	2.12(2.97)	*t*(52)=.40	-.11	-			
	Post-test	3.86(4.65)	3.46(4.12)	*t*(52)=.33	-.09	-	.14	21.51***	.01
Non-Word Reading	Pre-test	1.75(2.76)	1.65(2.70)	t(52)=.13	-.04	-			
	Post-test	3.79(3.87)	3.12(3.32)	*t*(52)=.68	-.19	-	.23	25.26***	.41

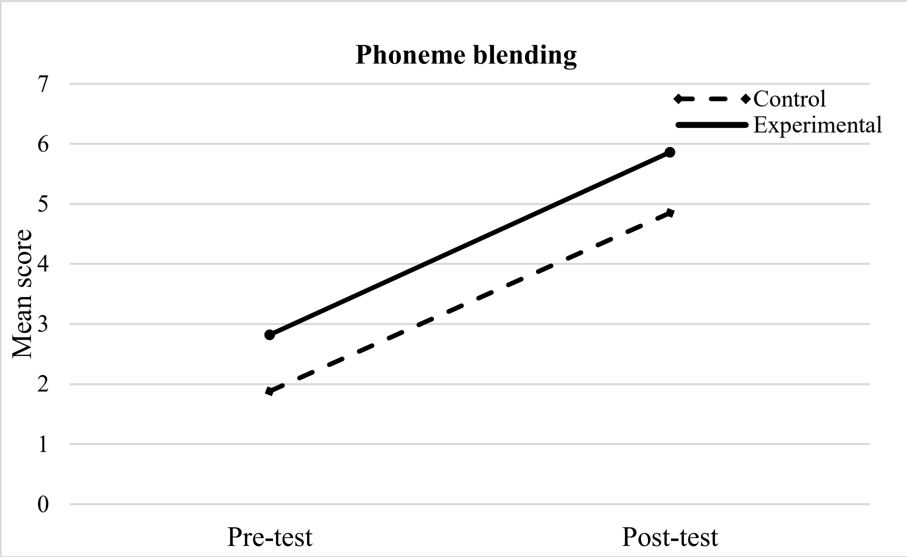
(Table continues)

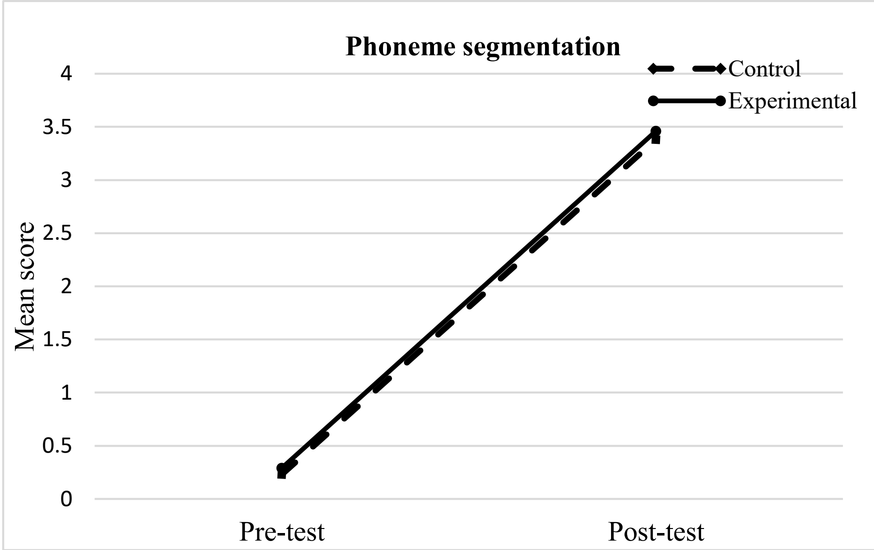
Measure	Assessment	Experimental n = 28 *M*(SD)	Control n = 26 *M*(SD)	*t*	*d*	*U*	Group effect *F*(1,52)	Time effect *F*(1,52)	Interaction effect *F*(1,52)
Composite	Pre-test	.05(.99)	-.06(.91)	*t*(52)=.42	.11	-			
Word reading	Post-test	.04(.99)	-.05(.89)	*t*(52)=.39	.11	-	.18	.000	.002
Oral reading	Pre-test	4.68(6.77)	3.58(5.75)	*t*(52)=.64	.18	-			
fluency	Post-test	8.96(10.79)	6.88(7.98)	*t*(52)=.80	.22	-	.57	35.17***	.58
PPVT-5	Pre-test	37(14.18)	42.85(9.29)	*t*(46.92)=1.80	.48	- -	-	-	-

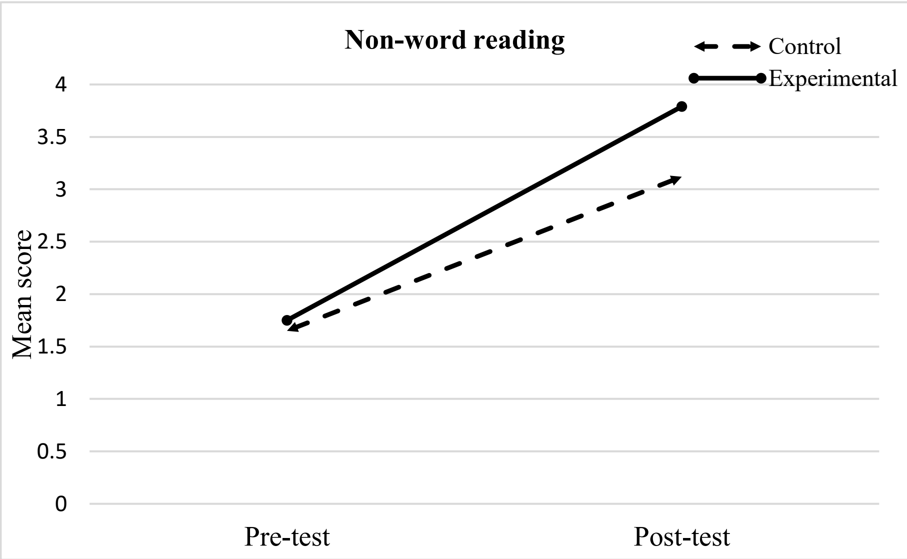
*p≤.05, **p≤.01, ***p≤ .001; [a]effect size calculated using Mann Whitney U test results

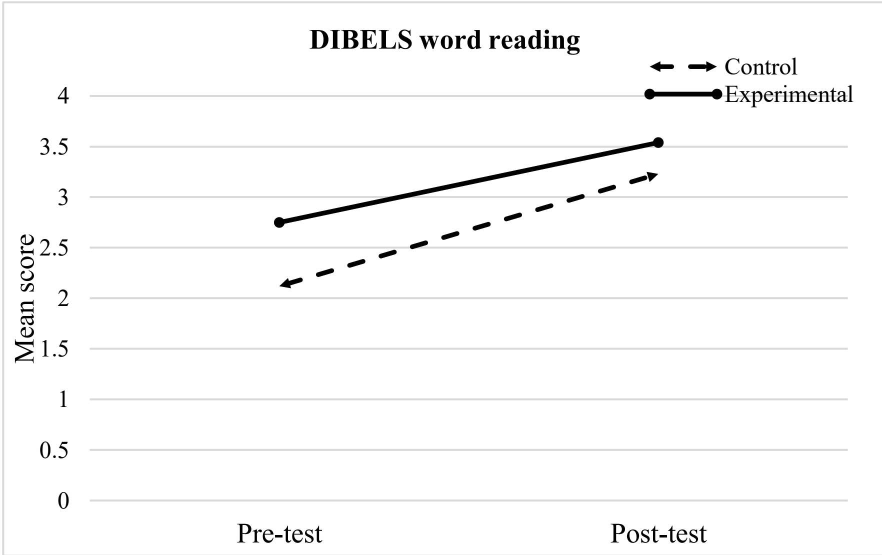
Figure 2 *Group comparisons of development from pre-test to post-test on oral-and-paper-based task*

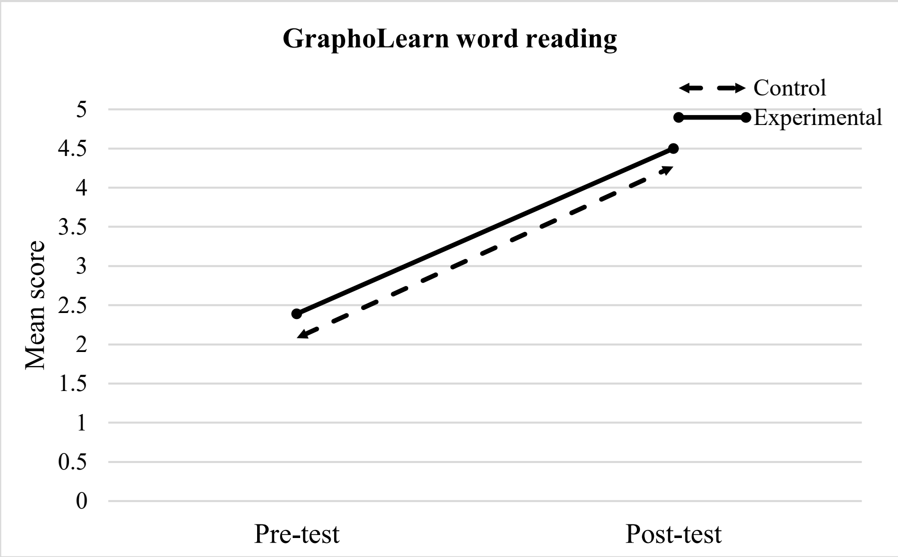


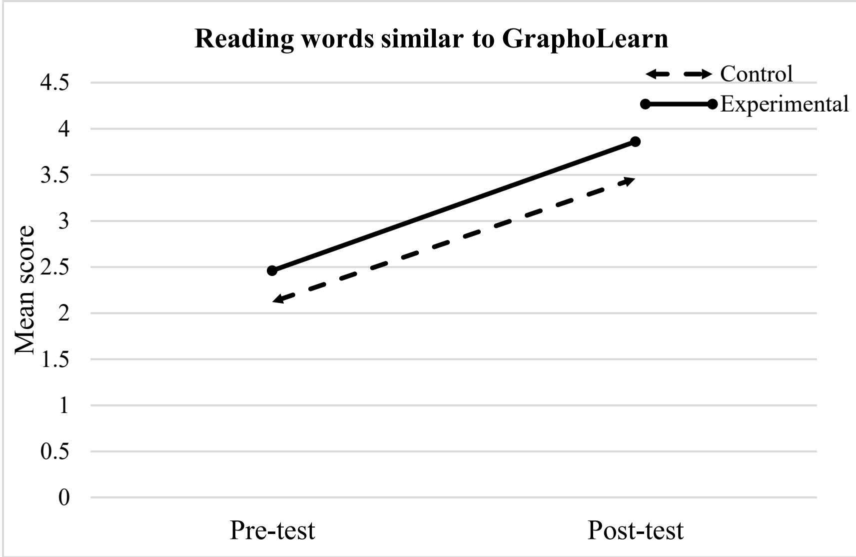

Phoneme blending
Control
Experimental
Mean score
0
1
2
3
4
5
6
7
Pre-test
Post-test

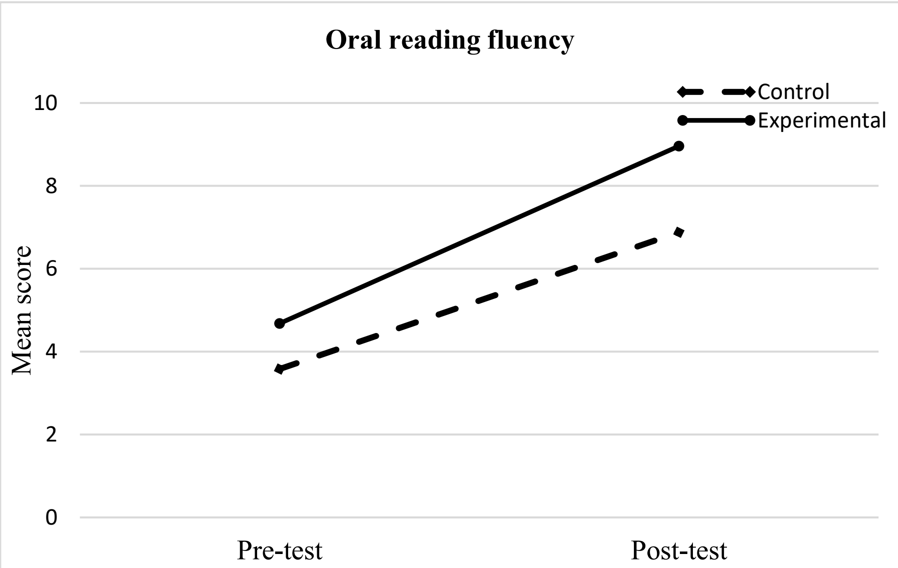

Phoneme segmentation
Control
Experimental
Mean score
0
0.5
1
1.5
2
2.5
3
3.5
4
Pre-test
Post-test


Non-word reading
Control
Experimental
Mean score
0
0.5
1
1.5
2
2.5
3
3.5
4
Pre-test
Post-test


DIBELS word reading
Control
Experimental
Mean score
0
0.5
1
1.5
2
2.5
3
3.5
4
Pre-test
Post-test


GraphoLearn word reading
Control
Experimental
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Mean score
Pre-test
Post-test


Reading words similar to GraphoLearn
Control
Experimental
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Mean score
Pre-test
Post-test


Oral reading fluency
Control
Experimental
10
8
6
4
2
0
Mean score
Pre-test
Post-test

4.3 Association Between GraphoLearn-Aligned Classroom Instruction and Post-test Scores

Finally, to estimate the potential role of students' attendance to the GraphoLearn aligned classroom instruction in predicting students' foundational literacy skills, correlation between attendance to the instruction and students' post-test scores in both in-game and oral-and-paper based measures were calculated. Pearson's correlations coefficients for association between all study variables are reported in Table 4. Also, Pearson's correlations coefficients for association between attendance to the GraphoLearn aligned classroom instruction and in-game and oral- and paper-based post-test measures are reported in table 5. In regard to the in-game measures (letter sounds, rime units and word recognition), the correlations were low and non-significant (r=.10-.25). In regard to the oral-and-paper based measures, there were significant low to moderate correlations between phonics instruction and letter name knowledge (r=.33, p=.01), letter sound knowledge (r=.41, p=.001), and phoneme blending (r=.35, p=.01). For the rest of the oral-and-paper tasks the correlations were non-significant. This led us to conclude that students' attendance to the GL-aligned classroom instruction is related to their letter name knowledge, letter sound knowledge and phoneme blending skill at post-test. Therefore, these significantly correlated variables were used for further analysis and the rest were excluded. Hierarchical regression analysis was conducted to examine the role of attendance to GraphoLearn-aligned classroom instruction as a predictor of oral-and-paper based letter name knowledge, letter sounds knowledge and phoneme blending.

In case of letter name knowledge, the results showed that attendance to the GraphoLearn-aligned classroom instruction do not predict its development in students (see Table 6). Group and vocabulary entered in Step one significantly predicted letter name knowledge at post-test (F(2,51)=3.39, p=.04) by explaining 12% of the variance. In this model, group differences were not statistically significant, indicating that being in the experimental or control group did not contribute to students' letter name knowledge. However, students' oral vocabulary skills were significantly predicting their letter-name knowledge. In the second step, when the pre-test score of letter-name knowledge was entered into the model, the explanation rate of the model increased statistically significantly by 78% (F(3,50)=139.85, p<.001). Such a significant contribution to the post-test scores indicated that students with better letter name knowledge at pre-test likely improved their knowledge at post-test. Also, probably they improved more than the students with poor letter name knowledge. After adding children's pre-test scores, vocabulary no longer had a significant effect on children's letter name knowledge. In the third step, the measure of students' attendance to the GL-aligned classroom instruction was added to the model. The total variance explained by the model was 91% ($F(4, 49)=109.47$, $p<.001$). Attendance to the GL-aligned classroom instruction explained 1% of variance over and above the control variables but remained statistically insignificant. Nevertheless, the result indicates that students who attended more GL-aligned classroom instruction, developed faster in the letter name knowledge. The pre-test was identified as the single strongest predictor but since the results were at ceiling, the result is not surprising ($\beta=.92$, $t=17.75$, $p<.001$).

In case of letter-sound knowledge, the results showed that attendance to the GraphoLearn-aligned classroom instruction uniquely and significantly predicted the development of letter sound knowledge (see Table 7). Group and vocabulary entered in Step one predicted letter sound knowledge at post-test ($F(2,51)=4.32$, $p<.05$) by explaining 15% of the variance. Here as well, the intervention groups did not make a significant contribution to the variance, and students' oral vocabulary skills predicted their letter-sound knowledge. Next, pre-test score of letter-sound knowledge entered in the second step increased the explanation rate of the model statistically significantly by 7% ($F(3,50)=4.52$, $p<.01$). In this step, both vocabulary and pre-test significantly predicted students' letter sound knowledge at post-test, indicating that students' oral vocabulary and letter sound knowledge benefited them through the intervention period and contributed to the gains at post-test. Thereafter, in the third step, attendance to GraphoLearn aligned classroom instruction was added to the model. The total variance explained by the model was 33% ($F(4,49)=5.94$, $p<.001$) wherein attendance to GraphoLearn aligned classroom instruction explained 11% of variance over and above contribution by group, vocabulary and pre-test score. After this addition, students' letter sound knowledge at pre-test no longer had a significant effect on their post-test gains. However, vocabulary continued to have a significant effect. Attendance to GraphoLearn aligned classroom instruction was thus found to be uniquely associated with letter sound knowledge ($\beta=.37$, $t=2.87$, $p<.01$). These results indicated that greater attendance to GraphoLearn-aligned classroom instruction improved students' letter-sound knowledge, thus underlying the efficacy of the attendance to the classroom instruction irrespective of the exposure to GraphoLearn.

In case of phoneme blending, again the results showed that GL-aligned classroom instruction uniquely and significantly predicted development in students' blending skills (see Table 8). Here as well, group and vocabulary added in Step one predicted phoneme blending skills at post-test ($F(2,51)=5.58$, $p<.01$) by explaining 18% of the variance. Like letter sound knowledge, only students' oral vocabulary skills made a significant contribution to the variance in their phoneme blending skills at post-test. Next, pre-test score of phoneme blending was entered into the model in Step two which increased the explanation rate of the model statistically significantly by 30% ($F(3,50)=15.05$, $p<.001$). This indicated that students with better pre-intervention phoneme blending skills developed more causing better learning gains at post-test than the students with poor pre-test score. This suggests that students' phoneme blending skills will support further development of the skill in a stable and continuous manner. In the third step, attendance to the GraphoLearn-aligned classroom instruction was added to the model. The total variance explained by the model was 53% wherein attendance to GraphoLearn-aligned classroom instruction explained 5% of variance over and above contribution by group, vocabulary, and pre-test score. Attendance to the GraphoLearn-aligned classroom instruction was thus found to be uniquely associated with phoneme blending ($\beta=.23$, $t=2.22$, $p<.05$). This means that students who attended GraphoLearn-aligned classroom instruction more, developed greater phoneme blending skills. Moreover, exposure to the instruction uniquely contributed to the development of blending skills over and above students' phoneme blending skill level at pre-test. Since students' pre-test score also significantly predicted their blending skills at post-test, it indicated that students with better phoneme blending skills improved further. Also, irrespective of one's initial skill level, students benefited from attending GraphoLearn-aligned classroom instruction.

To illustrate the learning gains in the above three measures related to the amount of exposure to the GraphoLearn-aligned classroom instruction, whole group's development from pre- to post-test in the three measures is reported using boxplots (see Figure 3). In case of letter name knowledge, the median score improved from 49 at pre-test to 49.50 at post-test. Since students were at ceiling at pre-test, there was not enough room for further development. However, between both pre- and post-test level, the variation in students' scores reduced and the outliers moved closer to the lower quartile of the whole group. In case of letter-sound knowledge, the whole group's median score improved from .00 to 27, thus showing an improvement of 27 points in students' learning gains on a total score of 54. These gains are quite steep given the average duration (~20 days) of students' attendance to GraphoLearn-aligned classroom instruction. However, there was a high amount of variation in students' letter sound knowledge, thus indicating that not all students benefited from the instruction equally. Even though students with scores at floor improved, the range of their improvement was wide. The gains in phoneme blending score showed left-skewed distribution, indicating many students' scores were at 50^{th} percentile. These students could probably improve more with more exposure to the instruction. Also, it could be possible that lack of letter sound knowledge at pre-test influenced the extent of learning gains students made by attending GraphoLearn-aligned classroom instruction.

Similar variation in students' scores was found in the case of phoneme blending measure as well. The whole intervention group's median score improved from 2 to 6 on a total score of 10, showing an improvement of 4 points. There was a variation in the development of phoneme blending skills. Some students' scores continued to be at floor and for some the scores improved from pre-test level. This could be because of students' pre-intervention phoneme blending skills. However, since significant contribution from students' attendance to GraphoLearn-aligned classroom instruction to the development of their phoneme-blending skills was also noted, it would be worthwhile to explore what limited the gains for some students. Inadequate pre-intervention blending skills or oral vocabulary level, both of which predicted students' scores, could be potential limitations.

Table 4 *Pearson's correlation for oral-and-paper based measures at pre-test and post-test level.*

	1	2#	3	4	5	6#	7	8	9	10	11	12	13	14	15	16
Pre-test measures																
1 Letter Name	-															
2 #Letter Sound	.17	-														
3 Initial Phoneme	.24	.56***	-													
4 Last Phoneme	.27	.32*	.37**	-												
5 Blending	.39	.56***	.35**	.59***	-											
6 #Segmenting	.17	.22	.32*	.25	.30*	-										
7 Word Reading	.45***	.48***	.31**	.34*	.56***	.29*	-									
8 Fluency	.42***	.46***	.21	.33**	.53***	.01	.87***	-								
9 PPVT	.37**	.10	.13	.24	.22	.05	.22	.19	-							
Post-test measures																
10 Letter Name	.95***	.32*	.33**	.30*	.46***	.14	.47***	.42**	.34*	-						
11 Letter Sound	.66***	.33*	.25	.14	.51***	.14	.39***	.37**	.32*	73***	-					
12 Initial Phoneme	.60***	.39**	.36**	.20	.55***	.11	.52***	.49***	.41**	.66***	.70***	-				
13 Last Phoneme	.63***	.37**	.32*	.25	.58***	.27*	.52***	.40***	.45***	.70***	.67***	.78***	-			
14 Blending	.68***	.52***	.39**	.41**	.65***	.31*	.61***	.53***	.35*	.78***	.71***	.61***	.69***	-		
15 Segmentation	.52***	.39**	.27	.32*	.72***	.28*	.59***	.50***	.32*	.57***	.68***	.66***	.71***	.68***	-	
16 Word Reading	.54***	.52***	.23	.32*	.61***	.24	.86***	.77***	.29*	.59***	.56***	.65***	.63***	.75***	.71***	-
17. Fluency	.46***	.52**	.18	.23	.48***	.18	.87***	.90***	.23	.46***	.47***	.53***	.46***	.59***	.53***	.80***

*p < 0.05 **p<0.01 ***<p<0.001; Note: # = Spearman's rho is reported for measures letter sound pre-test and segmenting pre-test

Note: Word reading indicated here is the composite word reading measure and Fluency is the oral reading fluency measure.

Table 5 *Pearson's correlation for GraphoLearn-aligned classroom instruction and post-test measures*

	GL-aligned classroom instruction
GL-aligned classroom instruction	-
In-game measures	
Letter-sound recognition	.25
Rime unit recognition	.25
Word recognition	.10
Oral-and-paper based measures	
Letter name knowledge	.33*
Letter sound knowledge	.41**
First phoneme identification	.22
Last phoneme identification	.24
Phoneme Blending	.35**
Phoneme Segmentation	.21
Composite word reading	.16
Fluency	.11

*p < 0.05 **p<0.01

Table 6 *Regression analysis summary on attendance to the GraphoLearn-aligned classroom instruction predicting letter names knowledge.*

	Step 1			Step 2			Step 3		
	Std. beta	*t*	*p*-value	Std. beta	*t*	*p*-value	Std. beta	t	*p*-value
Group (Control, Experimental)	.03	.25	.81	-.00	-.05	.96	-.02	-.47	.64
Vocabulary	.35	2.58	.01	-.01	-.14	.89	.01	.12	.91
Letter name Pre-test score				.95	19.10	<.001	.92	17.75	<.001
GL-aligned class instruction							.08	1.69	.10
ΔR^2		.12			.78			.01	

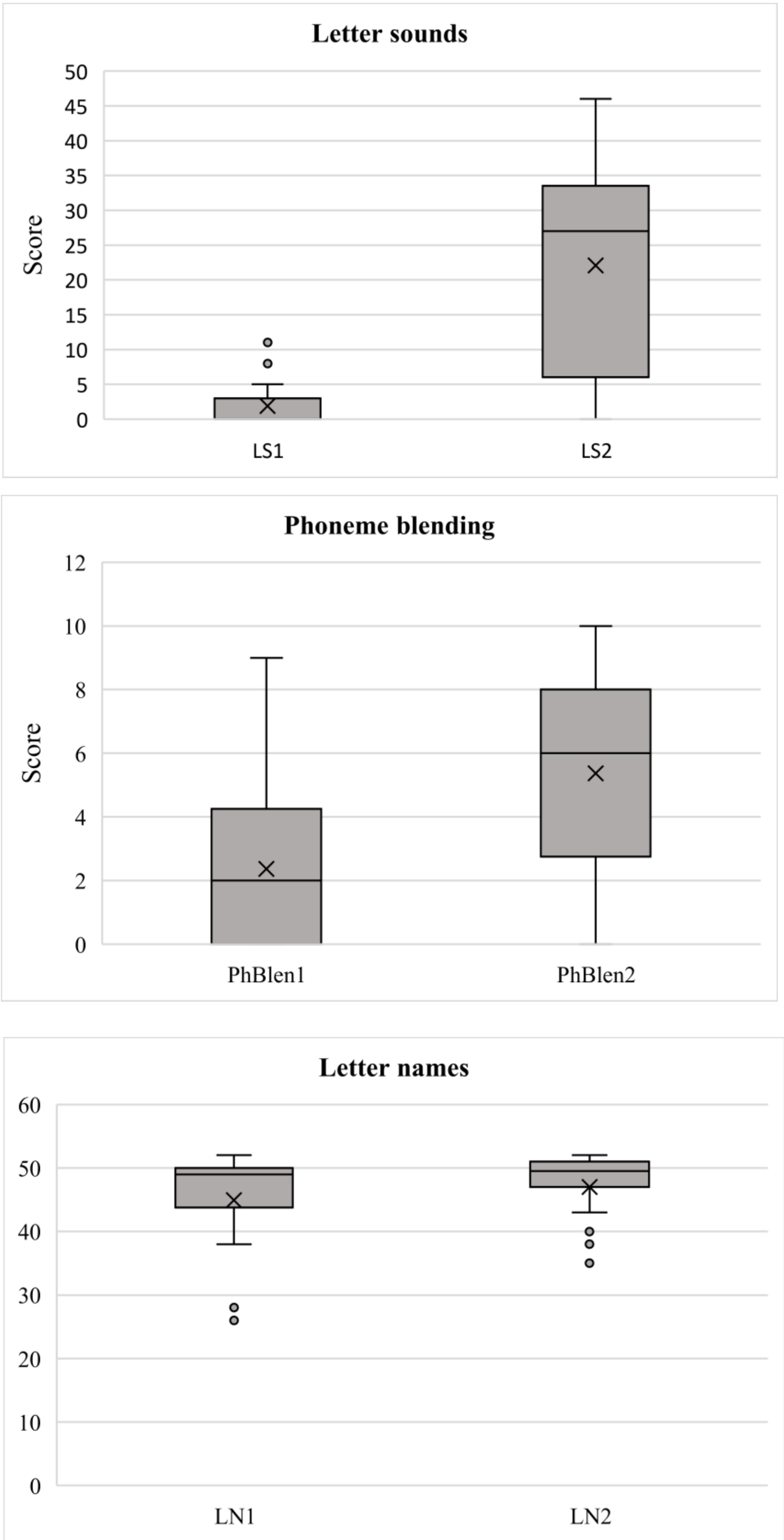
Table 7 *Regression analysis summary on attendance to the GraphoLearn-aligned classroom instruction predicting letter sounds knowledge.*

	Step 1			Step 2			Step 3		
	Std. beta	*t*	*p*-value	Std. beta	*t*	*p*-value	Std. beta	*t*	*p*-value
Group (Control, Experimental)	.22	1.62	.11	.16	1.18	.24	.09	.71	.48
Vocabulary	.37	2.77	.01	.22	2.59	.01	.36	2.97	.01
Letter Sound Pre-test score				.27	2.08	.04	.15	1.16	.25
GL-aligned class instruction							.37	2.87	.006
ΔR^2		.15			.07			.11	

Table 8 *Regression analysis summary on GraphoLearn- aligned classroom instruction predicting phoneme blending skills.*

	Step 1			Step 2			Step 3		
	Std. beta	*t*	*p*-value	Std. beta	*t*	*p*-value	Std. beta	*t*	*p*-value
Group (Control, Experimental)	.25	1.94	.06	.11	1.02	.31	.07	.61	.55
Vocabulary	.41	3.11	.00	.24	2.22	.03	.27	2.49	.02
Phoneme blending Pre-test score				.58	5.30	<.001	.52	4.88	<.001
GL-aligned class instruction							.23	2.22	.03
ΔR^2		.18			.30			.05	

Figure 3 *Development from pre-test to post-test in oral-and-paper based tasks for letter names, letter sound and phoneme blending for the whole group.*



LS1 – Letter sound pre-test

LS2 – Letter sound post test

PhBlen1 – Phoneme blending pre-test

PhBlen2 – Phoneme blending post-test

LN1 – Letter name pre-test

LN2 – Letter name post-test

5. DISCUSSION

The present study examined whether GraphoLearn combined with GraphoLearn-aligned classroom instruction could effectively support the development of foundational literacy skills of struggling readers in India. The participants were Grade 2 students who were learning English as a non-native language in an English-medium school in India. The study also examined how attendance to the GraphoLearn-aligned classroom instruction was related to the development of students' foundational literacy skills. Both the experimental and control group received 30 minutes of GraphoLearn-aligned classroom instruction 2-3 times a week. In addition, the experimental group played GraphoLearn, and the control group played Math Kids for 20 mins every day for 40 days over a period of 2.5 months. At the end of the intervention, all students had an average of 19.91 days of attendance to GraphoLearn-aligned classroom instruction. Both groups of students also played their respective games for an average of 10.26-11.18 hours. The intervention results showed that the experimental group had greater in-game learning gains over the control group but did not show transfer of learning to oral-and-paper based measures. Also, attendance to the GraphoLearn-aligned classroom instruction significantly contributed to the development of letter sound knowledge and phoneme blending skills of the whole class over and above experiment and control group membership, oral vocabulary level, and pre-test skills.

These results are critical for understanding the efficacy of the intervention, especially in the light of previous findings on implementing GraphoLearn in the context of India. Findings by Patel and her colleagues (2018, 2021) indicated that while GraphoLearn alone led to gains on in-game measures of letter-sound knowledge, rime-units knowledge and word recognition, it did not result in transfer of learning to a non-game-based environment. This finding has been consistent across various GraphoLearn studies in developed countries as well where classrooms are better resourced and where teachers provide instruction on phonics regularly (see McTigue et al., 2019 for review). Various studies (Patel et al., 2018, 2021; McTigue et al., 2019) have recommended that involvement of teachers can be a step towards translating in-game outcomes to out-of-game learning. The current study attempted to examine whether implementing GraphoLearn-aligned classroom instruction combined with GraphoLearn can facilitate transfer of skills learned in the game and lead to higher learning outcomes in oral-and-paper-based measures. However, despite making considerable gains on the in-game measures and exposure to the GraphoLearn-aligned classroom instruction, the experimental group's learning gains did not transfer to the oral-and-paper based measures. Nonetheless, the students' attendance to the GraphoLearn-aligned classroom instruction predicted learning gains in the oral-and-paper based letter-sound knowledge and phoneme blending. The results pointed at the efficacy of the instruction as students regardless of their group membership improved at post-test, thus underlying that providing only classroom instruction could work with or without GraphoLearn.

5.1 Effectiveness of GraphoLearn

Learning in a game becomes beneficial and meaningful only when students can apply it in a context outside the game (McTigue & Uppstad, 2018). However, in the current study, the learning gains made by the experimental group in GraphoLearn remained limited to the game environment itself. The group showed significantly higher post-test scores in all three in-game measures – GL letter sounds, GL rime units, and GL word recognition – compared to students who did not play GraphoLearn. In fact, they showed significantly faster development in the in-game skills indicating that their letter-sound knowledge developed quickly, and they could apply it to rime units and word recognition tasks effectively. These gains were observed despite students' poor literacy level. Prior to the intervention, students in both the groups could recognise an average of five out of 24 letter-sound correspondences. However, after the intervention, they identified 15 letter-sound correspondences, three times the pre-test level. Large effects of the exposure to GraphoLearn were observed in case of rime units and word recognition as well. However, development in these skills did not transfer in oral-and-paper based measures which could be due to reasons such as items used for assessment, duration of the intervention, students' early literacy level and inadequate GrapoLearn-aligned classroom instruction.

In this study, the out of game measures examined a range of skills to understand the transfer of learning. Thus, both sub-lexical skills (letter-name and letter-sound knowledge; phoneme identification; phoneme blending and segmentation) and lexical skills (word reading and oral reading fluency) were examined. Like previous studies (see McTigue et al., 2019 for review), the experimental group in this study also showed more gains in sub-lexical skills such as letter-sound knowledge than in the lexical skills such as word reading. These gains were tested using experimental measures instead of standardised measures to assess children on content that is closely aligned to GraphoLearn, thereby making them more sensitive to the treatment (Cheung & Slavin, 2013). Yet, the experimental group's learning gains in these measures were not significantly greater than the control group. In addition, transfer of learning was also examined by using items trained in GraphoLearn such as letter sounds, and word reading in the paper-based tasks. Use of trained items in paper-based letter-sound knowledge and GraphoLearn word reading measures dialled up their relevance to the experimental group as it has been noted that trained items show greater transfer than untrained items (Görgen et al., 2020). However, in the study the effects of the intervention were not in the favour of the experimental group in oral-and-paper based tasks for trained and untrained items.

Lack of significant gains even on the trained items was a surprising finding since based on previous studies (Patel et al., 2022; McTigue et al., 2019) it was believed that students exposed to GraphoLearn combined with GraphoLearn-aligned classroom instruction could show transfer of in-game skills to paper medium. Nevertheless, comparatively larger effect size observed in the case of letter sounds and phoneme blending paper-based measures suggest that students developed more in measures that assessed them on items closer to the content taught throughout the intervention. Since there was no GraphoLearn only group in this intervention, it is not possible to point out the cause(s) of the observed effect. To better understand the effectiveness of more attendance to the GraphoLearn-aligned classroom instruction in facilitating transfer, it would be valuable to examine the effects of exposure to only GraphoLearn game.

Development and transfer of skills could be influenced by students' duration of their exposure to GL and the extent in-game progress (Patel et al., 2021) as the game requires players to score a minimum of 80% to move ahead in the game. This results in practice tailored to a student's level as well as variation in children's progress. In a study examining transfer of in-game skills to out-of-game context, Patel and her colleagues (2021) found that students with better pre-existing skills made more in-game progress and eventually showed more gains in the in-game learning. Also, transfer of these gains was observed in pseudoword reading, a measure that shows specific underlying phonological awareness in students (Nag-Arulmani et al. 2003). Another study (Ahmed et al., 2020) examining the benefits of more in-game progress noted that the GL Rime group who've crossed over 50% of streams (Stream 16, level 5) made significant progress in phonics decoding skills and reading in English.

These findings are meaningful because learners in the current study had very less pre-existing decoding skills and completed an average of 11.5 streams. Only three out of 28 students from the experimental group reached stream 25 and 57% of the group crossed stream 8 from where instruction and practice in rimes units and word formation become explicit (Patel et al., 2021). Since the GraphoLearn content across the 25 streams was incorporated into all the 26 sessions of the GL-alignedclassroom instruction, variations in students' game progress meant that even though a student was stuck at Stream 1 throughout the intervention period, that student had to attend the classroom instruction on sound units from other streams as well. Thus, students received exposure to different letter sounds at the same point of time which could have been confusing for them. Moreover, since only 50% of students crossed stream 11, level 5 in GraphoLearn, these students learnt new sounds and orthographic rime units only during the GL-aligned classroom instruction and could not practice all the sound units in GraphoLearn. The experimental group probably could have shown a significant growth had both GraphoLearn-aligned classroom instruction and in-game instruction focused on same set of sound units throughout (Mo et al., 2016). The efficacy of such forms of combined instruction must be explored to examine how GraphoLearn and GraphoLearn-aligned classroom instruction can build on each other's potential.

While the classroom instruction fell short of aligning itself to students' in-game learning level, it provided them the opportunity to verbalise their learning, thus creating a room for practising sound units along with the teachers. Prior studies have shown that in-game learning remains 'inert' or situated within the game (see McTigue et al., 2019 for review), or that it remains 'intuitive' (McTigue & Upstad, 2019) without such supports. McTigue & Upstad indicated that even though a game is designed for interactivity, students do not necessarily engage with it mindfully. Similar observations had been reported by Ecochard (2015) in the GraphoLearn study in Peru. Lack of mindful engagement probably got more worse due to the language of instruction – English – which students were unable to understand. However, since the experimental group showed in-game learning gains despite the language barrier, the opportunity to verbalise their learning along with teachers became much more relevant for these students as they could produce and practice pronunciation of these sounds. Yet, these learning gains remained limited to the game-environment only. These findings bring forth fundamental questions on using computer-assisted learning tools for supporting struggling readers' foundational literacy skills.

5.2 Role of GraphoLearn-aligned Classroom Instruction

Despite the poor early literacy levels observed in students, improvement in all the in-game and oral- and paper-based measures was observed for both the experimental and control group. Amongst these measures, the improvement in the oral- and paper-based letter-sound knowledge and phoneme blending was uniquely predicted by students' attendance to GraphoLearn-aligned classroom instruction. These findings reaffirm the potential role of consistent, systematic and explicit classroom instruction on phonics in the development of literacy skills with or without GraphoLearn. While exposure to GraphoLearn combined with GraphoLearn-aligned classroom instruction was beneficial for the experimental group's in-game literacy skills, it did not contribute to students' outside-the-game skills. In fact, the results of the current study indicated that GraphoLearn-aligned classroom instruction significantly predicted students' letter sound knowledge and phoneme blending skills. This indicated a gap in the added value of GraphoLearn in the development of students' foundational literacy skills. Of course, since phoneme blending is not explicitly taught in GraphoLearn (Patel et al., 2021), it could be unfair to expect the tool to support the development of blending skills.

These results emphasised the efficacy of the amount of GraphoLearn-aligned classroom instruction combined with GraphoLearn in supporting struggling readers in a resource constrained learning environment. It was surprising to note students' improved learning levels in such a short duration of intervention time. At pre-test, the experimental and control group had a mean score of 2.43 and 1.31 respectively on a total of 52 in the letter-sound paper measure. Post the classroom instruction, these scores increased to 23.96 and 20.08 respectively. Similar floor effects and post-intervention gains were seen in initial and last phoneme identification and phoneme blending measures. Such a development in students' learning scores point out the promise of improving student learning level irrespective of their initial literacy level. The intervention seemed to support both low and high scoring students, however we cannot say which group of students benefited or would benefit more from the teacher only or computer only or a combined form of instruction. These gains observed for Grade 2 students in the intervention duration raises questions on the effectiveness of pre-intervention classroom instruction that students received. Despite the exposure to English medium instruction in Grade 1, students' pre-intervention learning level is poor. However, with only an average instructional time of approximately 20 days, students' learning gains are quite considerable. In word reading measures for instance, students' reading level is still low but their knowledge of words and non-word reading skills have doubled. These findings show the value in using systematic and explicit classroom instruction for building decoding skills and thereby, preventing reading failure (Saine et al., 2011). Furthermore, they reaffirm the efficacy of implementing instruction through teachers (McEwan, 2017).

Attendance to GraphoLearn-aligned classroom instruction predicted gains in measures where there was immense room for growth, unlike paper-based letter-name knowledge where students were already at ceiling. Both letter sound knowledge and phoneme blending had immense room for growth and the hierarchical regression analysis showed that students who attended more GraphoLearn-aligned classroom instruction improved more than those who did not. Moreover, this improvement showed a wide individual variation in students learning level, indicating that some students improved, and some did not. These findings brings us to the question why the attendance to the classroom instruction did not significantly predict gains in other oral and paper-based measures which had a room for growth. One reason could be that the instruction did not focus on phoneme segmentation and word reading. Thereby, no significant relation was observed between the attendance to the GraphoLearn-aligned classroom instruction and segmentation and word reading skills. In the case of initial and last phoneme identification also, explicit instruction on phoneme identification was not provided as the lesson plans on GraphoLearn-aligned classroom instruction closely followed the content and its sequence within GraphoLearn.

Second reason could be the difficulty level of the skill measured. Phoneme segmentation is a complex skill, as identifying single units, phonemes, in spoken words is more difficult than identifying onsets and rimes (Ehri, 2022; Stahl & Murray, 1998) and developmentally precedes blending skills (Holopainen et al., 2020). Also, as phoneme identification is a more difficult skill as opposed to phonemes blending and rimes identification, its acquisition takes time as English language readers mostly rely on larger sub-lexical units (Zeigler & Goswami, 2005). Since average attendance to the GraphoLearn-aligned classroom instruction was less (only ~19.91 days), it could be a potential cause for the lack of significant relation between phoneme identification skills and GraphoLearn-alignd classroom instruction. Nevertheless, these lexical and sub-lexical skills at post-test showed correlation with the attendance to the instruction, and with each other. Moreover, the effect size of group differences observed in case of initial and last phoneme identification, phoneme segmentation, and word reading measures also emphasised that GraphoLearn combined with GraphoLearn-aligned classroom instruction had a small effect on these measures. Third reason could be the nature of the measure itself. Each of these measures were more distal to the treatment than the letter sounds and phoneme blending measures. Thus, we did not see strong effects of the group differences on the distal measures. It is likely that with more attendance to the instruction, students' phonological awareness and reading skills could be further developed.

5.3 Barriers and Opportunities for Classroom Instruction on Phonics

The findings on the development of students' letter sounds knowledge and phoneme blending skills bring forth a possible role of using students' first language, Hindi in this case, in the development of their English literacy skills. In the case of the GraphoLearn-aligned classroom instruction, the language of instruction could have been an advantage in learning, as teachers provided instruction in Hindi – students L1, a language all students could understand at least orally. This could have allowed for an incidental use of L1 (Hindi) for supporting English decoding through shared phonological awareness skills, a relationship that has been identified in previous studies (Koda, 2008; Piasta & Hudson, 2022) Given that GraphoLearn does not provide instruction in students' native language, they might have faced difficulty in understanding the in-game instruction (Kim et al., 2016), thereby being unable to utilise it fully, to the extent of building transferable skills.

The relationship between Hindi and English is not as apparent as it may look. English is an alphabetic language, while Hindi is an alpha-syllabary writing system. Thus, using Hindi in teaching letter-sounds in English calls for a careful consideration of parallels that could be drawn between different sound units in both languages. Teachers in the current study had a mistaken understanding of sounds; they used English language phonemes interchangeably with Hindi *aksharas*, which are either vowels or consonant-vowel syllables. For example, sound /m/ in English was mistaken as sound म - /mə/, leading to a confusion about the sounds and letters that correspond with each other. Since this misunderstanding was shared by both teachers and students as well, the classroom instruction required unlearning prior knowledge of sounds and their corresponding letter names. However, as mentioned earlier, despite the differences between the phonological and orthographic aspects of Hindi and English, students' phonological awareness in Hindi could have incidentally supported them in developing decoding skills in English (Koda et al., 2008, Patel et al., 2022).

Use of Hindi sound units for English phonics presents an opportunity more than a predicament. Utilising Hindi as a resource for teaching English language could in fact make the instruction more effective as it could draw upon students' knowledge of sounds in Hindi to teach sounds in English language (Ball, 2011; Ludwig et al., 2019), thereby preventing students from using both set of sounds interchangeably. This finding finds more strength from a study (Nishanimut et al., 2013) which used students' L1 - Kannada, another alpha-syllabary language for teaching English (L2) letter sounds and found that students' phonemic skills developed better when they were built over their existing metalinguistic knowledge, i.e., Kannada *aksharas* than when they were only built through teaching L2 phonemes. The *aksharas* could be deconstructed into units representing phonemes and thus be used in an alphabetic manner for teaching English sound units.

Irrespective of the method used for the instruction on letter-sounds, it is important to utter correct pronunciation of phonemes. During classroom observations it was noted that teachers and students struggled with letters that have multiple pronunciations and with phonemes that had multiple spellings. For instance, /*o*/ has different pronunciation when used as a long vowel (as in 'old') and short vowel (as in 'hot') – a concept with which teachers were unfamiliar. Teachers also found it difficult to pronounce phonemes such as /*y*/ and /*t*/ correctly. They also shared their confusion regarding similarities between phonemes /*a*/, /*e*/, and /*r*/, /*l*/ in GL and during classroom instruction. Thus, despite exposure to the game, provisions for use of resources such as picture books and flashcards, and access to scripted lesson plans, teachers at times struggled with the pronunciation of phonemes.

Challenges with phonics classroom instruction were also observed in the number of phonemes covered in a lesson and students' response to the same. 5-7 letter sounds and orthographic rimes were taken up in a 30-minute lesson including 2-3 previously taught sounds as opposed to 1-3 target sounds in a week (Vadasy and Sanders, 2021). Due to this, even though the instruction was persistent and progressive (Piasta & Hudson, 2022), it was difficult to provide more frequent practice distributed over a period of time. More and consistent exposure to words with same orthographic rimes could have helped students to read new words with same rime pattern correctly (Conrad, 2009; Jones & Reutzel, 2012; Zeigler & Goswami, 2005). Due to more number of sound units per lesson, repetition of phonemes and rimes were probably not adequate and thereby limited acquisition of sound units and their identification during basic decoding (Sunde, Furnes & Lundetræ, 2020). In addition to limited practice and repetition, it was noted that students' attention span dropped after 15-20 minutes of instruction, and they required an energiser to refocus their attention on the instruction. This could be because too many letter sounds were introduced in a lesson. Exposure to more new content requiring multiple cognitive processing could lead to 'heavy paired associate learning' which could be confusing and frustrating for early learners and eventually lead to disengagement (Hulme et al., 2007). Vadasy and Sanders (2021) have noted that struggling young readers benefited from exposure to three correspondences per week. These findings are meaningful as they indicate that equal amounts of instruction for all sound units may not be suitable (Jones & Reutzel, 2012) and appropriate pacing of the units could improve students' learning experience and outcomes. (Vadsay & Sanders, 2021).

The challenges with instruction were further magnified by the class size and implementation of the classroom instruction to the whole class. During the instruction two grade 2 level classes of 40 students sat together in their classroom due to lack of infrastructure in the school. Teachers struggle to instruct a class of 80 students and follow their learning levels was evident during classroom observations and was also shared as a concern post the classroom instruction. Teaching a large class size is difficult as teachers cannot cater to all students with varying achievement levels (Duflo et al. 2011). This context of the intervention is in stark contrast to the studies showing efficacy of computer assisted reading instruction (CARI) combined with teacher instruction in small groups (Up to 25) or as a remedial lesson outside the regular class hours (Cheung & Slavin, 2013; Saine et al., 2011). This contrast brings forth the potential of integrating GraphoLearn into classroom instruction in the context of high teacher-student ratio – a reality of many classrooms in India. This reality also comprises of inadequate learning environment, stunted pre-literacy skills, ineffective classroom teaching and lack of teachers skilled in phonics instruction.

Within the above context, it is imperative to examine which characteristics of computer-assisted learning such as GraphoLearn and GraphoLearn aligned classroom instruction can support each other and predict gains in foundational literacy in English language. This examination is pertinent in the light of findings from previous studies indicating that exposure to EdTech alone is ineffective in building foundational literacy skills outside the game (Mc Tigue et al., 2019; Rodriguez-Segura., 2021) and findings from this study showing that the exposure to the classroom instruction predicts foundational literacy over and above exposure to GraphoLearn. To examine the opportunities that GraphoLearn can afford when integrated into classroom instruction, findings pertaining to the first research question will be now discussed.

5.4 Piecing Together the Classroom Instruction and Education Technology

The above findings indicate the challenges and opportunities in integrating GraphoLearn into classroom instruction and raise questions on the added value that the in-game instruction brings to students' learning. In the absence of transfer of learning despite the exposure to the classroom instruction, the opportunity cost for the experimental group seems to be high. Despite the additional instruction time on GraphoLearn and the in-game learning gains, the group did not show significant growth over the control group outside the game. To address this concern, it is important to consider two important questions – how an adaptive learning software, GL in this case, must be implemented to maximise the potential of the tool; and who is likely to benefit most from exposure to the game?

The promise of adaptive tools lies in their potential to provide personalised learning thereby preventing struggling learners from falling behind. Muralidharan and his colleagues (2019) indicate that when an adaptive learning tool shows content based on students' assessed learning level and adjusts their exposure to content based on their progress, then the system can accommodate varying learning levels as well as students' pace of learning. An average classroom in public schools in India show a range of 4 grade levels of learning achievements in language within the class, and this range goes on to increase up to 6 grade levels in higher grades (Muralidharan et al., 2019). Teaching a class with such a wide dispersion of learning levels in a resource-constrained environment could be a challenge even for well trained teachers. In this context, GraphoLearn can provide personalised learning thereby increasing students' access to the learning opportunity which remains inaccessible due to limited differentiation and delayed feedback in the classroom instruction by teachers. However, for students at risk of reading difficulties, playing the game even as a supplement to a rote-and repetition-based instruction by teachers may not develop students' phonological awareness. The students in fact may get frustrated if they are unable to move ahead with the streams unless the teacher intervenes to understand where students need help and provides that. Thus, if the underlying approach to literacy instruction is inadequate, then a standalone GL or a similar EdTech intervention is unlikely to result in significant improvement in learning achievement (Kim et al., 2016; Conn, 2017; McTigue et al., 2019).

It is important that students are able to transfer their skills to reading, spelling and writing (Ehri, 2022), otherwise, there is no point of learning games (Bainbridge et al., 2022; McTigue & Upstad, 2018). For transfer in other contexts, mindful interaction with the game is must (Bainbridge et al., 2022). A teacher could support that by modelling to students how they can slow down while playing and think about their choices before clicking any option (Saldaña, 2013). To support students with the in-game learning, the teacher must also know what instruction to provide and how. Since many teachers lack the necessary content knowledge on the development of foundational literacy skills (Kim et al., 2016), a tool like GL can be a starting place for them as well for learning phonics. However, to be able to provide effective classroom instruction, a teacher must also know how to organise and provide phonics instruction. Only availability of GL will not necessarily bring a change in teachers' behaviour and skill level (Piper et al., 2016). To address this, scripted lesson plans on systematic phonics instruction can be a preliminary step towards providing a high-quality literacy instruction (Kim et al., 2016; Piper et al., 2016). This would also support teachers in ensuring that there is no misalignment between in-game and in-class instruction.

Teachers' role in making adaptive technologies more effective has not been explored at length because tools such as GL are often used to provide instructional support outside of regular classroom instruction (Major et al., 2020; Miglai & Burch, 2019). The findings of this study reinforce the need to examine the efficacy of alternative interventions integrating technology into classroom instruction such as the one in the current study. A profound increase in students' sub-lexical skills, even in the face of a less than optimal learning environment and extremely high teacher-student ratio and despite the presence of technology is indicative of the indispensable role of teachers in supporting struggling readers (He et al., 2008). Prior studies have shown that the efficacy of self-paced instruction and classroom instruction varies for low and high performing students or students at or below their grade level (Muralidharan et al., 2019; see Rodriguez-Segura, 2021 for review). Higher performing students have been observed to benefit more from self-paced intervention (Mo et al., 2016) thereby indicating better suitability of tools like GL for students at a higher learning level. Students with a learning level at floor are likely to get stuck at the initial levels (Patel et al., 2021), as observed in this intervention as well. Struggling readers can in fact benefit more from systematic and explicit phonics-based instruction (He et al., 2008; Sunde et al., 2020) as they are less likely to show transfer of skills (Cheung & Slavin, 2013). These findings indicate that the efficacy of an intervention with both technology and classroom instruction component must be examined for these high and low-performing sub-group of learners to identify how computer-assisted classroom instruction can be tailored for students at opposite ends of the learning spectrum.

5.5 Limitations

While this study makes an important contribution to the efficacy studies on GraphoLearn and studies in the Indian context, its efficacy is bound by its limitations. Lack of control group, i.e. the group without classroom instruction prevented the study from examining the transfer of learning from within- to outside-the-game measures. Also, the GraphoLearn-aligned classroom instruction time was limited due to which 6-7 sound units were introduced to children in each class. This was contrary to the recommendations by previous studies on teaching appropriate number of sound units, three letters a week (Vadsay & Sanders, 2021). Teaching of sound units was also limited by teachers' professional capacity as they first learned to produce the correct sound units and then taught the students. Teachers themselves did not get enough practice and sometimes pronounced incorrect sound units during the classroom instruction. The researcher was present throughout to support them with the instruction and clarifying their doubts before and after every GraphoLearn-aligned classroom instruction session. While this ensured high fidelity of the intervention, it limited perspective on how teachers would be supported if this study is implemented at scale.

Since this intervention was a small RCT, its results are local and may be limited in their relevance for other contexts of the developing countries. The study involved a small-scale intervention with a small sample size and within a single context, thereby limiting generalisability of its findings. Thus, it is important to have similar efficacy studies in other contexts and larger sample size as well towards identifying who, what and how can GL support. In addition, the presence of a researcher throughout the intervention, even though a strength, was a limitation from a practical lens. The researcher was present in every GraphoLearn-aligned classroom instruction session full-time to ensure the fidelity of the intervention and followed up with teachers to check their understanding of phonics. This may not be possible under practical conditions, especially if this design is implemented at scale where there would be more implementors and exercising such fidelity would be difficult. Also, the intervention in this study required multiple gadgets; procuring and using them may not be feasible in public schools as they may not have access to the gadgets.

5.6 Practical implications

Overall, this study made a unique contribution to the role of GraphoLearn combined with GraphoLearn-aligned classroom instruction in supporting students' phonological awareness and reading skills. It is notable that even with challenging learning conditions and little support to teachers on phonological awareness and phonics instruction, GraphoLearn-aligned classroom instruction contributed to the gains in students' early literacy skills. Irrespective of students' exposure to the GraphoLearn and their skills at pre-test, the GraphoLearn-aligned classroom instruction from teachers benefited students, thus indicating the merit in getting more exposure to the instruction. At present, there is no research study in the Indian context that examines the role of teachers in the light of EdTech interventions and explores how they perceive use of EdTech in classrooms (Miglani & Burch, 2018). The findings of this intervention are a step towards bridging this gap by providing insights into how to integrate GraphoLearn into regular classroom instruction to support foundational English language literacy skills of struggling readers in English-medium government schools. These findings are also relevant in the light of the recommendations on developing phonological awareness to support children's foundational literacy skills (NCERT, 2022).

The current study enforces the need for improving classroom instruction even in technology interventions and highlights the role of teachers. In the GraphoLearn-aligned classroom instruction, implemented for a mean exposure time of 19.91 days, students were taught six to seven sound units in each lesson. This instruction could be planned better in future studies keeping in mind how students using GraphoLearn can be supported through the greater attendance to the classroom instruction covering less sound units in a week. Since children with low pre-reading skills require more time to acquire letter-sound correspondences and learn to decode and read (Saine et al., 2011), incorporating a gradual introduction to a new set of sounds can be beneficial for students (Vadsay & Sanders, 2021). Thus, future studies should focus on providing classroom instruction with a smaller number of sound units and a greater number of implementation days to determine whether that helps students with transfer of in-game skills to the oral-and-paper medium.

GraphoLearn can be an effective tool for practising phonics, however its efficacy needs further examination to determine how skills developed in the game can be transferred from game. Towards this, bridging the instructional content in the classroom and in the game can be an approach (Bainbridge et al., 2022). Also, it is important to explicitly tell students that GraphoLearn competency is the goal, otherwise playing for fun will remain the default for students (McTigue & Upstad, 2018). McTigue & Upstad (2018) suggest that before introducing students to a game like GraphoLearn, teachers must introduce the target concept using hands on tools such as picture cards, and only after students have demonstrated a required level of competence, the teacher should introduce the game to students. Thus, in this context, GraphoLearn can be used as a tool for practice and reinforcement. Moreover, it can be a powerhouse of game analytics and provide insights into players' behaviours. GraphoLearn can be better integrated into the lesson plans if the game analytics are readily accessible and understandable by teachers or researchers. This could also make classroom interaction more meaningful for students, thereby raising the effectiveness of adult interaction (McTigue et al., 2019) and aiding transfer of learning.

For greater transfer effects, GraphoLearn could integrate in-game instruction or subtitles in students' native language, Hindi which has more than 422 million native speakers according to the 2001 census (Office of the Registrar General & Census Commissioner, 2001). Instruction in a familiar language would be beneficial for students as it has been noted that when students are provided instruction in an unfamiliar language, they skip the parts they do not understand (Central Square Foundation, 2019). The relevance of using L1 is also underlined by studies on cross-linguistic transfer indicating that students' phonological and reading skills in L1 transfer during L2 reading acquisition (Ludwig et al., 2019; Patel et al., 2022). Thus, GraphoLearn-aligned classroom instruction can include lesson plans which use L1 as a resource and thereby bridge gaps in learners' letter-sound knowledge and purge the misconceptions on parallels between Hindi and English sound units.

Furthermore, to be able to provide effective literacy instruction, teachers must understand phonological awareness and reading skills along with how to provide instruction for the development of these skills (Cunningham & Stanovich, 2004). Thus, teachers can utilise GraphoLearn to support their phonics skills which should also be supported through professional development (PD) on phonics and its instruction. By playing the game, teachers can connect their instruction and students' game experience, thereby embedding GraphoLearn time into the formal classroom instruction. Since students' phonological awareness skills are being built through the classroom instruction, it is important that teachers pronounce letter sounds correctly. Thus, PD support towards phonics skills development could be provided as well.

To sum, the study, combining an adaptive game and classroom instruction on foundational literacy, is a step towards understanding instructional approaches that could be suitable and effective for large classroom environments in a developing country. The findings also make an important contribution to randomised control trial evaluations of technology-based English literacy interventions in India.

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

ASER Centre (2014). ASER 2005-2014 reading tasks. http://img.asercentre.org/docs/Aser%20survey/Call%20for%20papers/assessments_1a.pdf

ASER Centre (2022). Annual status of education report (rural) 2022. https://img.asercentre.org/docs/ASER%202022%20report%20pdfs/All%20India%20documents/aserreport2022.pdf

Bainbrige, K., Shute, V., Rahimi, S., Liu, Z., Slater, S, Baker, R, S., & D'Mello, S, K. (2022). Does embedding learning supports enhance transfer during game-based learning? *Learning and Instruction*, *77*. https://doi.org/10.1016/j.learninstruc.2021.101547

Ball, J. (2011). Enhancing learning of children from diverse language backgrounds: Mother tongue-based bilingual or multilingual education in the early years. Paris: UNESCO. http://unesdoc.unesco.org/images/0021/002122/212270e.pdf

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

Banerjee, A. & Duflo, E. (2012). *Poor economics: A radical rethinking of the way to fight global poverty*. New York, NY: Public Affairs

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., & Walton, M. (2016). Mainstreaming an effective intervention: Evidence from randomised evaluation of 'teaching at the right level' in India. Working paper 22746. National Bureau of Economic Research. http://www.nber.org/papers/w22746

Bhattacharjea, S., Wadhwa, W., & Banerjee, R. (2011). Inside primary schools: A study of teaching and learning in rural India. ASER Centre, New Delhi.

Bold, T., Filmer, D, P., Martin, G., Molina, E., Rockmore, C., Stacy B, W., Svensson, J., & Wane, W. (2017). What do teachers know and do? Does it matter? Evidence from primary schools in Africa. The World Bank Policy Research Working Paper Series 7956, pp. 2–35. https://openknowledge.worldbank.org/server/api/core/bitstreams/513f39b9-1249-58d7-8de4-dae2a5614171/content

Boyer, N., & Ehri, L. C. (2011). Contribution of phonemic segmentation instruction with letters and articulation pictures to word reading and spelling in beginners. *Scientific Studies of Reading*, *15*(5), 440– 470. https://doi.org/10.1080/10888 438.2010.520778

Buckingham, J. (2020). Systematic phonics instruction belongs in evidence-based reading programs: A response to Bowers. *The Educational and Developmental Psychologist*, *37*, 105–113. https://doi.org/10.1017/edp.2020.12

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/10.1177/15291 00618 772271

Central Square Foundation. (2019). The EdTech Lab Series: Insights from rapid evaluations of EdTech products. IDInsight. https://www.idinsight.org/wp-content/uploads/2021/09/EdTech-Lab-Report_November-2019.pdf

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. K. & 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

Clark, D.B., Tanner-Smith, E.E., & Killingsworth, S.S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79–122. https://doi.org/10.3102/0034654315582065

Cohen, R.A., Mather, N., Schneider, D.A., & White, J.M. (2017). A comparison of schools: Teacher knowledge of explicit code-based reading instruction. *Reading and Writing*, *30*, 653– 690. https://doi.org/10.1007/s11145-016-9694-0

Conn, K. M. (2017). Identifying effective education interventions in Sub-Saharan Africa: A meta-analysis of impact evaluations. *Review of Educational Research*, *87*(5), 862-898. https://doi.org/10.3102/0034654317712025

Conrad, N. J. & Levy, B. A. (2009). Training letter and orthographic pattern recognition in children with slow naming speed. *Reading and Writing*, *24*(1), 91–115. https://doi.org/10.1007/s11145-009-9202-x

Cunningham, A. E., & Stanovich, K. E. (2004). Disciplinary knowledge of K-3 teachers and their knowledge calibration in the domain of early literacy. *Annals of Dyslexia*, *54*(1), 139–167. https://doi.org/10.1007/s11881-004-0007-y

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. https://doi.org/10.1080/09243453.2011.625125

Duflo, E., Pasacaline, D., & Kremer, M. (2011). Peer effects, teacher incentives, and the Impact of Tracking: Evidence from a randomized evaluation in Kenya. *The American Economic Review*, *101*(5), 1739–1774. DOI:10.1257/aer.101.5.1739

Dunn, D. M. (2019). *Picture Peabody Vocabulary Test* (5th ed.) [Measurement instrument]. Bloomington, MN: NCS Pearson.

Dutta, U., & Bala, N. (2012). Teaching of English at primary 4level in government schools. *National Council of Education Research and Training*. http://www.ncert.nic.in/departments/nie/del/publication/pdf/English_Primary_level.pdf

Ecochard, S. (2015). Learning to read with GraphoGame, an ethnography in Peruvian rural school. A post-graduate thesis in Education. Jyväskylä: University of Jyväskylä.

Ehri, L, C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Ehri, L, C. (2020). The science of learning to read words: a case for systematic phonics instruction. *Reading Research Quarterly*, *55*(S1), S45–S60. https://doi.org/10.1002/rrq.334

Ehri, L, C. (2022). What teachers need to know and do to teach letter– sounds, phonemic awareness, word reading, and phonics. *The Reading Teacher*, *76*(1), 53–61. https://doi.org/10.1002/trtr.2095

Escueta, M., Nickow, A. J., Oreopoulos, P. & Quan, V. (2020). Upgrading education with technology: Insights from experimental research. J*ournal of Economic Literature*, *58*(4), 897–996. https://doi.org/10.1257/jel.20191507

Folotiya et al., 2014 - Jere-Folotiya, J., Chansa-Kabali, T., Munachaka, J.C., Sampa, F., Yalukanda, C., Westerholm, J., & 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), 41–436. https://doi.org/10.1007/s11423-014-9342-9

Glewwe, P., Hanushek, E, A., Humpage, S. & Ravina, R. (2013). School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010. In Paul Glewwe (Eds.), *Education Policy in Developing Countries* (pp. 13–64). University of Chicago Press.

Glewwe, P. & Muralidharan, K. (2016). Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications. In Eric Hanushek, Stephen Machin and Ludger Woessmann (Eds.), *Handbook of the Economics of Education* (pp. 653–744). Elsevier.

Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Gonzalez- Frey, S. M., & Ehri, L. C. (2021). Connected phonation is more effective than segmented phonation for teaching beginning readers to decode unfamiliar words. *Scientific Studies of Reading*, *25*(3), 272–285. https://doi.org/10.1080/10888 438.2020.1776290

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

He, F., Linden, L, L., & MacLeod, M. (2008). How to teach English in India: Testing the relative productivity of instruction methods within the Pratham English language education program. Unpublished manuscript, Columbia University, New York, http://www.leighlinden.com/PicTalk%20Working%20Paper%202008-07-02.pdf

Holopainen, L., Koch, A., Hakkarainen, A., & Kofler, D. (2020). Predictors of reading skills at the first and second grade: The role of orthography. *Reading Psychology*, *41*(5), 461–484. https://doi.org/10.1080/02702711.2020.1768988

Hudson, K, H., Moore, K, A., Han, B., Koh, P, W., Binks-Cantrell, E., & Joshi, R. M. (2021). Elementary teachers' knowledge of foundational literacy skills: a critical piece of the puzzle in the science of reading. *Reading Research Quarterly*, *56*(S1), S287–S315. https://doi.org/10.1002/rrq.408

Hulme, C., Goetz, K., Gooch, D., Adams, J., & Snowling, M. J. (2007). Paired-associate learning, phoneme awareness and learning to read. *Journal of Experimental Child Psychology*. *96*(2), 150–166. https://doi.org/10.1016/j.jecp.2006.09.002

Jones, C. D. & Reutzel, D. R. (2012). Enhanced alphabet knowledge instruction: Exploring a change of frequency, focus, and distributed cycles of review. *Reading Psychology*, *33*(5), 484-464. https://doi.org/10.1080/02702711.2010.545260

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies:

The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Advances in psychology: Orthography, phonology, morphology, and meaning* Vol. 94. (pp. 67–84). Amsterdam, Netherlands: North-Holland

Kim, Y.-S. G., Boyle, H. N., Zuilkowski, S. S., & Nakamura, P. (2016). Landscape Report on Early Grade Literacy. Washington, D.C.: USAID. https://allchildrenreading.org/wp-content/uploads/2019/07/USAID-Landscape-Report-on-Early-Grade-Literacy.pdf

Koda, K. (2008). Impacts of prior literacy experience on second language learning to read. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second language literacy development* (pp. 68–96). Routledge

Kyle et al., 2013 - Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., and Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions, GG Rime and GG Phoneme. *Reading Research Quarterly*. *48*, 61–76. doi: 10.1002/rrq.038

Linden, L. (2008). Complement or substitute? The effect of technology on student achievement in India. Working paper 44863. https://poverty-action.org/sites/default/files/publications/Gyan_Shala_CAL_2008-06-03.pdf

Lucas, T. & Villegas, A, M. (2010). A framework for preparing linguistically responsive teachers. In Tamara Lucas (Eds.) *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators*. Taylor & Francis Group.

Ludwig, C., Guo, K., & Georgiou, G. K. (2019). Are reading interventions for English language learners effective? A meta-analysis. *Journal of Learning Disabilities*, *52*(3), 220–231. https://doi.org/10.1177/0022219419 825855

Major, L., Francis, G. A. & Tsapali, M. (2020) The effectiveness of technology-supported personalised learning in low- and middle-income countries: A meta-analysis. *British Journal of Educational Technology*. *52*, 1935–1964. https://doi.org/10.1111/bjet.13116

McEwam, P, J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, *85*(3), 353–394. https://doi.org/10.3102/0034654314553127

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. & Upstad, P. H. (2018) 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

Menon, S., Kutty, S. S., Apte, N., Basargekar, A., & Krishnamurthy, R. (2019). Powerful Reforms in Early Language and Literacy Instruction in India. In Nic Spaull & Jon Comings (Eds.) *Improving Early Literacy Outcomes: Curriculum, Teaching, and Assessment* (pp. 191-206). Brill Sense.

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/10.1177/09731849188031

Mo, D., Bai, Y., Boswell, M. & Rozelle, S. (2016). Evaluating the effectiveness of computers as tutors in China: Impact Evaluation Report 41. International Initiative for Impact Evaluation.

Muralidharan, K. and Singh, A. (2020). Improving public sector management at scale? Experimental evidence on school governance in India. RISE Working Paper Series. 20/056. https://doi.org/10.35489/BSG-RISE-WP_2020/056

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*, 187–219. https://doi.org/10.1111/1467-9922.00153

Nag, S. (2013). Low literacy attainments in school and approaches to diagnosis: an exploratory study. *Contemporary Education Dialogue*, *10*(2) 197-221. https://doi.org/10.1177/0973184913484997

Nag-Arulmani, S., Reddy, V & Buckley, S. (2003). Targeting phonological representations can help in the early stages of reading in a non-dominant language. *Journal of Research in Reading*, *26*(1), 49-68. https://doi.org/10.1111/1467-9817.261005

NCERT (2021a). National Achievement Survey. New Delhi: National Council of Educational Research and Training. https://nas.gov.in/report-card/nas-2021

NCERT (2021b). National Initiative for Proficiency in Reading with Understanding and Numeracy. Department of School Education and Literacy. Government of India. https://ncert.nic.in/pdf/announcement/NIPUNBharat_web.pdf

NCERT (2022). National Curriculum Framework for Foundational Stage. https://ncert.nic.in/pdf/NCF_for_Foundational_Stage_20_October_2022.pdf

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

Office of the Registrar General & Census Commissioner. (2001). Population Census. https://censusindia.gov.in/nada/index.php/catalog/20028

Ojanen, E., Ronimus, M., Ahonen, T., Chansa-Kabali, T., February, P., Jere- Folotiya, J., et al. (2015). GraphoGame-a catalyst for multi-level promotion of literacy in diverse contexts. *Frontiers in Psychology*, *6*(671). https://doi.org/10.3389/fpsyg.2015.00671

Patel, P., Torppa, M., Aro, M., Richardson, U., and 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

Patel et al. 2022 - 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*, *38*(1), 76-89. https://doi.org/10.1111/jcal.12592

PEW. (2021). China has a larger middle class than India, where most people are low income. Pew Research Center.

https://www.pewresearch.org/fact-tank/2021/03/18/in-the-pandemic-indias-middle-class-shrinks-and-poverty-spreads-while-china-sees-smaller-changes/ft_2021-03-18_indiachina_02/

Piasta, S. B., Purpura, D. J., & Wagner, R. K. (2010). Fostering alphabet knowledge development: A comparison of two instructional approaches. *Reading & Writing*, *23*, 607– 626. https://doi.org/10.1007/s11145-009-9174-x

Piasta, S. B. & Hudson, A. K. (2022). Key knowledge to support phonological awareness and phonics instruction. *The Reading Teacher*, *76*(2), 201-210. doi:10.1002/trtr.2093

Piper, B., Zuilkowski, S, S., Kwayumba, D. & Strigel, C. (2016). Does technology improve reading outcomes? Comparing the effectiveness and cost-effectiveness of ICT interventions for early grade reading in Kenya. *International Journal of Educational Development*, *49*, 204-214. https://doi.org/10.1016/j.ijedudev.2016.03.006

Pitchford, N, J. (2015). Development of early mathematical skills with a tablet intervention: A randomised control trial in Malawi. *Frontiers in Psychology*, 6. https://doi.org/10.3389/fpsyg.2015.00485

Pittman, R.T., Zhang, S., Binks Cantrell, E., Hudson, A., & Joshi, R.M. (2020). Teachers' knowledge about language constructs related to literacy skills and student achievement in low socio-economic status schools. *Dyslexia*, *26*(2), 200– 219. https://doi.org/10.1002/dys.1628

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of Game-Based Learning, *Educational Psychologist*, *50*(4), 258-283. https://doi.org/10.1080/00461520.2015.1122533

Rice, M., Erbeli, F., Thompson, C, G., Sallese, M, R., & Fogarty, M. (2022). Phonemic awareness: a meta-analysis for planning effective instruction. *Reading Research Quarterly*, *57*(4), 1259-1289. https://doi.org/10.1002/rrq.473

Richards- Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners. *Exceptional Children*, *82*(2), 144–169. https://doi.org/10.1177/00144 02915585483

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.201405 281859

Rodriguez-Segura, D. (2022). EdTech in Developing Countries: A Review of the Evidence. *World Bank Research Observer*, *37*(2), 171-203. https://doi.org/10.1093/wbro/lkab011

RTE. (2009). Right to Education Act 2009. Ministry of Education. Government of India https://www.education.gov.in/en/sites/upload_files/mhrd/files/upload_document/RTE_Section_wise_rationale_rev_0.pdf

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.ezproxy.jyu.fi/10.1111/j.1467-8624.2011.01580.x

Sankar, D. & Linden, T. (2014). How much and what kind of teaching is there in elementary education in India? Evidence from three states. South Asia Human Development Sector Report No. 67. Washington, DC: The World Bank.

Sinha, S. (2019). Early literacy instruction in India. In Nic Spaull & Jon Comings (Eds.) *Improving Early Literacy Outcomes: Curriculum, Teaching, and Assessment*. Brill Sense.

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

Stahl, S, A. & Murray, B. (1998). Issues involved in defining phonological awareness and its relation to early reading. In Jamie L. Metsala, Linnea C. Ehri (Eds.) *Word recognition in beginning literacy*. Taylor & Francis Group

Sunde, K., Furnes, B. & Lundetræ, K. (2020). Does introducing the letters faster boost the development of children's letter knowledge, Word Reading and Spelling in the First Year of School?, *Scientific Studies of Reading*, *24*(2), 141–158. https://doi.org/10.1080/10888438.2019.1615491

Vadsay, P.F & Sanders, E. A. (2021). Introducing grapheme-phoneme correspondences (GPCs): exploring rate and complexity in phonics instruction. *Reading and Writing*, *34*(1), 109–138. https://doi.org/10.1007/s11145-020-10064-y

Vagh, (2012). Validating the ASER Testing tools: comparison with reading fluency measures and the Read India measures. ASER Centre. https://img.asercentre.org/docs/Aser%20survey/Tools%20validating_the_aser_testing_tools__oct_2012__2.pdf

Washburn, E.K., Joshi, R.M., & Binks-Cantrell, E.S. (2011). Teacher knowledge of basic language concepts and dyslexia. *Dyslexia*, *17*(2), 165–183. https://doi.org/10.1002/dys.426

World Bank. (2016). What is happening inside classrooms in Indian secondary schools? A time on task study in Madhya Pradesh and Tamil Nadu. The World Bank, Washington D.C.

World Bank Group. (2019). Learning poverty. The World Bank. https://www.worldbank.org/en/topic/education/brief/learning-poverty

Wouters, P. & Oostendorp, H. V. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computer and education*, *60*, 412–425. http://dx.doi.org/10.1016/j.compedu.2012.07.018

Ziegler, J.C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

APPENDICES

Appendix 1

JYVÄSKYLÄN YLIOPISTO

UNIVERSITY OF JYVÄSKYLÄ

CONSENT TO PARTICIPATE IN SCIENTIFIC RESEARCH

Name of the study: Testing the effectiveness of a technology-based reading intervention for children learning to read English in India
I understand that participation in the study is voluntary and that me and my child can stop participating at any time, without giving a reason. There will be no negative consequences for us if my child and I withdraw. The data collected about me and my child up to the point of withdrawal may still be used in the study.

I have been adequately informed about the study and the processing of my and my child's personal data. I have received the information sheet about the study, as well as the privacy notice. I have also had the opportunity to ask the researchers further questions.

I consent to provide the following data:
☐ Information in parent questionnaires
☐ Child's test results from paper-pencil tests and in-game tests

I confirm that my child and I will not participate in face-to-face data collection if we have flu symptoms, fever, are recovering from illness, or are feeling otherwise unwell.
Yes ☐ No ☐

I understand the information that I have received about the study and my child's participation and agree to participate in this study.
Yes ☐ No ☐

By signing this form, I accept that
- data will be collected from me, and my child as described in information sheet,
- my and my child's data can be used in accordance with the procedures outlined in the privacy notice.

I give my consent to the sections specified above by ticking the "yes" boxes.

If I or my child do not wish to participate in a particular section, I or my child have the right to refuse by ticking the "no" box.

However, I, on behalf of my child, still agree to participate in the study otherwise.

Confirmation
Signature:
Name:
Date:
Contact details:
Deepti Bora
M.A. Student,
University of Jyväskylä
+91-9821782147
bdeepti@student.jyu.fi

Appendix 2
UNIVERSITY OF JYVÄSKYLÄ

FACULTY OF EDUCATION
AND PSYCHOLOGY

Date: 31.05.2022

A description of the processing of personal data for scientific research purposes (privacy notice; Articles 13, 14 and 30 of Regulation (EU) 2016/679)

1. **Personal data processed in study - Testing the effectiveness of a technology-based reading intervention for children learning to read English in India**

The scientific goal of the study is to assess the effectiveness of an app-based tool – GraphoLearn – for developing English reading skills in students in India when complemented with classroom instruction from trained teachers. This intervention will be carried out in Grade 2 or 3 classrooms of an English medium government school or a low-budget private school. To effectively assess the learning gains and to understand the factors influencing the learning levels, we need information that could help us determine home literacy environment, exposure to English language inside and outside schools, education level of parents and teachers, and results from English tests and games.

The following personal data will be collected from you: Name, telephone number, educational qualifications, income level, number of children, mother language, first language, data log from game, questionnaire on demographic and home learning environment, English test results of your child, questionnaires on teacher knowledge.

This privacy notice will be given to the participants of the study in hard copy.

2. **Legal grounds for the processing of personal data for research/archiving purposes**

☒ Processing is necessary for scientific or historical research purposes or statistical purposes, and it is correctly proportional in relation to the goal in accordance with public interest (section 4.1(3) of the Finnish data protection act)

Transferring personal data outside the EU/EEA/India

During this study, your personal data will be transferred from India to Finland for the purposes of research using secure means.

Protection of personal data

In this study, the processing of personal data is based on a proper research plan, and a responsible person has been appointed for the study. Your personal data will only be used and disclosed for purposes of conducting historical or scientific research or for other similar purposes (statistics), and it is otherwise ensured that no data about you is disclosed to unauthorised parties.

Prevention of identifiability

☒ Data will be anonymised when it is generated (all identifiers will be fully removed so that no persons can be identified from the data, and no new data can be merged with the data)

Personal data used in the study will be protected by means of

☒ username ☒ password ☐ registered use ☐ access control (physical facilities)

☐ other, please specify:

Also, consent forms collected from the parents will be stored and processed securely. Since the consent forms have to be transported to Finland, they will be scanned and sent over NextCloud or a similar storage and sharing service. Data from questionnaires will be logged in and stored in a .csv file.

The processing of personal data *after* the study

☒ The research register will be anonymised, i.e. all identifiers will be fully removed so that no persons can be identified from the data, and no new data can be merged with the data.

Controller(s) and researchers

The controller is the party which, alone or with another party, defines the goals and means of the processing or personal data, as well as the organisation(s) and person(s), and is responsible for the lawfulness of processing.

The controller for this study is the r**esearcher - Deepti Bora**

Contact information:

bdeepti@student.jyu.fi
C/O Minna Torppa
RUU A226.1 Ruusupuisto
PO Box 35
FI-40014 University of Jyväskylä

Person in charge and the contact person of the study:

Deepti Bora
+358-40******** / +91-98*******
bdeepti@student.jyu.fi

Researchers:

Deepti Bora Master's Students Faculty of Education and Psychology University of Jyväskylä	Minna Torppa, PhD Professor Faculty of Education and Psychology, University of Jyväskylä	Priyanka Patel Doctoral Student Faculty of Education and Psychology University of Jyväskylä

Rights of data subjects

Right to access data (Article 15, GDPR)

You have the right to obtain information about whether your personal data is processed, and which personal data is processed. If required, you can request a copy of the personal data processed.

Right to have data rectified (Article 16, GDPR)

If there are any inaccuracies or errors in the processing of your personal data, you have the right to request your personal data to be rectified or supplemented.

Right to have data erased (Article 17, GDPR)

You have the right to request your personal data to be erased in certain situations. However, the right to have data erased does not exist if the erasure prevents the purpose of processing from being fulfilled for scientific research purposes or makes it much more difficult.

Right to the restriction of processing (Article 18, GDPR)

You have the right to restrict the processing of your personal data in certain situations, such as if you deny the accuracy of your personal data.

Right to object (Article 21, GDPR)

You have the right to object to the processing of your personal data if processing is based on public or legitimate interest. As a result, the university cannot process your personal data unless it can prove that processing is based on a significantly important and justified reason which supersedes your rights.

Derogation from the rights of data subjects

Derogation from the aforementioned rights is possible in certain individual situations on the basis of the GDPR and the Finnish data protection act, insofar as the rights prevent scientific or historical research purposes or statistical purposes being fulfilled or make it much more difficult. The need for derogation must always be assessed separately in each situation.

Profiling and automated decision making

In this study, your personal data will not be used in automated decision making. In this study, the purpose of the processing of personal data is not to assess your personal characteristics, i.e. profiling. Instead, your personal data and characteristics will be assessed from the perspective of broader scientific research.

Executing the rights of data subjects

If you have any questions about the rights of data subjects, please contact the university's **data protection officer**. All requests related to the execution of rights must be sent to the registry office of the University of Jyväskylä. Registry office and archive, P.O. Box 35 (C), 40014 University of Jyväskylä, tel.: +358 (0)40 805 3472, email: kirjaamo@jyu.fi. Visiting address: Seminaarinkatu 15, Building C (Main Building, 1st floor), Room C 140.

Any data breaches or suspicions of data breaches must be reported to the University of Jyväskylä.

https://www.jyu.fi/en/university/privacy-notice/report-data-security-breach

You have the right to file a complaint with the supervisory authority of your permanent place of residence or employment if you consider that the processing of personal data is in breach of the GDPR. In Finland, the supervisory authority is the Office of the Data Protection Ombudsman.

Contact for Office of the Data Protection Ombudsman: https://tietosuoja.fi/en/home

Appendix 3

RESEARCH NOTIFICATION

University of Jyväskylä

FACULTY OF EDUCATION
AND PSYCHOLOGY

Date: 31.05.2022

Name of study and controller

Name of the study: Testing the effectiveness of a technology-based reading intervention
for children learning to read English in India
Controller of the study: Deepti Bora

Request to participate in a study.

You and your child are requested to participate in the study regarding gains in early English language reading skills such as phonological awareness and decoding. In this intervention, the child will learn from a teacher who will be trained in early reading skills and provided lesson plans on the curriculum to be used in her classroom. After the teacher instruction, the child will play either GraphoLearn or an age-appropriate math game for 20minutes a day 4-6 days a week for a specified period of time during their regular school day, after which we will assess their progress on basic skills of English reading. GraphoLearn is a computer-based reading tool which provides training on the connections between spoken and written language through explicit instruction on letter-sound correspondences, a strong predictor of later reading skill. The main aim of this research is to determine whether GraphoLearn along with teacher instruction on phonics is effective at improving the English reading skills of children in English medium schools in India.

You are requested to participate in the study because your child's classroom has been selected for this study as your and your child's mother tongue and first language is not English and your child is studying in Grade 2 or Grade 3 in an English medium government school or a low-budget private school. This notification describes the study and participation in it.

Participating in the study requires that your mother tongue and first language is not English and that you send your child to an English or a Hindi-medium government or a low-budget private school in India. In total, minimum 80 research subjects will be requested to participate.

A survey will be conducted with parents to identify their educational qualifications and the literacy environment they provide to their children at home. Also, data will be collected from teachers. They will be pre- and post-tested on their knowledge of phonics. They will also be asked about their educational qualifications. Observation of classroom instruction will be conducted by the researcher on a bi-weekly basis.

Voluntariness

Participating in this study is voluntary. You can refuse to participate in this study or cancel your participation at any time. Providing personal data is not required on statutory or contractual grounds, or on the grounds of concluding a contract. Not providing the data does not have any consequences for you.

Progress of the study

Duration of the study: 11 weeks

Duration of responding to the survey and tests: 10 days (July 1 – 9; and September 5 – 9)

Number of research visits: 55 (The researcher will visit the school during school hours. The classroom instruction and GraphoLearn game time will then be provided to the child.)

Duration of research visits: July 1 – September 9

Content of research visits: The child will receive classroom instruction on phonics from a trained teacher trained for 15 mins. The teacher will be provided lesson plans to teach phonics content. After that the child will play a game Math game or an English game, GraphoLearn on a phone which will be provided by the researcher. GraphoLearn is a research-based game build for developing early readying skills.

Actions to be carried out:

1. Survey of parents and teachers
2. Teacher training and classroom observations
3. Pre- and post-assessment of research subjects
4. Classroom instruction on phonics lessons by teachers
5. Phonics instruction using GraphoLearn app or Math game

Any harm and discomfort resulting from the study

The study will not cause any harm. Also, the study does not pose any risks.

Research costs

No fee will be paid for participating in the study.
The study will be partly funded by University of Jyvaskyla, Finland.

Research results and their announcement

The results of the study will be reported in scientific publications, theses, congress and seminar presentations, and lectures. They may be used also in the development of practical applications.

Insurance coverage of research subjects

The staff and activities of the University of Jyväskylä are covered by insurance. The coverage includes insurance against treatment injury, liability insurance and voluntary accident insurance. During the study, research participants (test persons) are insured against accidents, damages and injuries caused by an external cause. Accident insurance is valid during physical tests and journeys *immediately related to the research. In addition to accidents, the insurance covers muscle or tendon sprains that are the direct result of a specific one-time exertion and movement and for which medical care has been delivered within 14 days from the injury. Compensation will be paid for a period that covers, at the most, six weeks from the date of the injury. Surgical operations and magnetic resonance imaging are not compensated for as treatment for a sprain caused by exertion and movement.*

Contact details for obtaining additional information

Deepti Bora
Master's Student
Faculty of Education and Psychology
University of Jyväskylä
bdeepti@student.jyu.fi
+358-********* / +91-982*******

Appendix 4 GraphoLearn based lesson plan

Lesson Plan 4	
Topic	Sounds and letters
Week 1	July 18 – 23
Time required	30 mins
Objectives	To recognise the names and sounds of letters *r, v, y, j, g, e, n, m, a, I, t* To recognise rime units using letters To read words and recognise phonemes and rimes used. To create rime units and words using letters.
Core concepts	Phoneme identification, phoneme replacement, rime identification
Target letters and sounds	*/r/, /v/, /y/, /j/, /g/, /e/, /n/, /m/, /a/, /i/, /t/*
Target rimes	*/im/, /it/, /in/*
Resources needed:	GraphoLearn app, flash cards, board and chalk
Pre-requirement	Teacher should have played GraphoLearn up till stream 3.
Tone setting	Set expectations for the day – What students are going to learn How students can learn best from the lesson Write topic on the board
Brief	In this session, we will be introducing five new letter sounds and one new rime. Also, we will review sounds done in earlier sessions.
Go back to go forward	Review previously done letter sounds and rimes Suggested activity: You may review the sounds in the class by – 1. Asking students to share the sound in a given rime or a common word 2. Asking students to share the sound for a given letter 3. Playing the audio clips of the sounds and asking students to repeat it 4. Showing pictures associated with a keyword and asking students to share a sound
Guided Inquiry	As you revise the sounds */i/, /t/, /m/,* and */n/*, form rime explicitly by showing blending of */i/* and */t/, /i/* and */m/,* and */i/* and */n/.* You can show the blending by modelling how sounds are combined. Model by saying iiiitttt = it and write on the board rime */it/*. Similarly, model formation of /im/ and /in/. Ask the students to repeat together /i/ /t/ = /it/. Similarly, other rimes. Introduce the sound */r/* by showing the flashcard for r/R. Ask students to identify the letter. Share the sound */r/* and then give an example of a word with */r/* sound. (For instance – red, rat). You may ask students to identify the letter sound in the middle and end of the word as well. Underline the le tter in the word one by one when asking the sounds. OR ask students to come to the board and underline the correct letter symbol for a given sound in the word. You may model blending of letter sounds */r/, /e/* and */d/* for students by slowly saying rrrreeedddd = red. Run your index finger through the letters while sharing their speech sounds to show the direction of reading and process of blending.

	Ask students to repeat this action. Students can come on the board and move their index finger in the same way as you. This routine can be set and repeated 4-5 times. Now, introduce the sound */v/*. Show the flash card and ask them to identify the letter name. Also, show a picture associated with a keyword containing the speech sound for */v/*. Practice speaking the sound by asking students to repeat words with */v/* sound such as voice, van, vest, love. Now introduce the sound /y/. Show the flash card and asking them to identify the letter name for both y and Y. Share the sound by playing the audio file and then say the speech sound aloud yourself. Ask the class to repeat the sound after you. Practice speaking the sound by asking students to repeat words with */y/* sound such as yellow, yes, yum. Now, repeat sounds again - /r/, /y/, /v/. You may ask one student at a time and then ask the whole class. Next, introduce the sound /j/ in the same way as above. Practice speaking the sound by asking students to repeat words with /j/ sound such as jug, juice, jump, jam.
Wrap Up	Wrap up – Revise the sounds and the blends once more. Suggested activity – Ask students to identify first sound in words starting with letter sounds they learnt today. You may also ask students – “What did you learn today?”

Appendix 5 – Letter names and Letter sounds task

t	n	f	y
I	R	D	G
Y	V	r	b
P	Z	i	c
A	O	J	x
h	K	o	S
M	q	U	w
v	a	F	u
C	m	L	d
N	X	e	W
g	B	E	j
H	l	s	Q
p	k	z	T

Appendix 6 – Initial phoneme identification task

ago	hot	for
wall	ride	pain
seen	year	mean
beach		

Last phoneme identification task

big	gas	shop
done	date	king
other	could	warm
week		

Appendix 7 – Phoneme blending task

/c/ /a/ /t/	/d/ /i/ /d/
/m/ /o/ /p/	/j/ /a/ /r/
/p/ /u/ /t/	/t/ /e/ /n/
/l/ /o/ /g/	/w/ /a/ /g/
/s/ /i/ /p/	/g/ /u/ /m/

Phoneme segmentation task

bit	sad
pot	gap
bug	hen
mix	pig
job	sun

Appendix 8: Word reading tasks

DIBELS word reading task

an	by	no
out	lot	for
jo	take	from
care	next	none
turn	place	speak
drive	sound	voice
could	earth	

GraphoLearn word reading task

sat	pin	tie
toe	mop	hall
ship	rag	long
card	right	clock
join	farm	good
soil	nurse	knife
discount	jacket	

Words similar to GraphoLearn reading task

bat	tin	pie
foe	top	ball
chip	bag	song
hard	light	frock
coin	harm	wood
boil	purse	knock
miscount	packet	

Non-word reading task

ib	op	com
nem	cug	sim
fet	hap	yot
nud	nirk	thon
nasp	kort	mame
phad	knent	chish
twint	sming	

Appendix 9: Oral reading fluency task

A big tree stood in a garden. It was alone and lonely. One day a bird came and sat on it. The bird held a seed in its beak. It dropped the seed near the tree. A small plant grew there. Soon there was another tree. The big tree was happy.