Bridging the gap in Brazilian literacy:

Kalulu phonics provides gains in 1st grade reading

Abstract

Purpose: A significant portion of Brazilian students fail to meet basic reading standards. The 2017 National Common Curricular Base (NCCB) marked a pivotal effort by the government to prioritize to recognize phonics in early reading instruction. However, the NCCB lacked clear directives for implementation, resulting in limited adoption in classrooms where teachers were accustomed to constructivist approaches to literacy. To address this shortcoming, we adapted the open-source Kalulu phonics method to Brazilian Portuguese and tested its effectiveness as a supplementary learning workstation during classroom reading instruction. The Kalulu method is available at XXX[1]. **Method:** Participated 184 Brazilian children (97 girls) from 1st year, aged 6 to 8 years (M = 6.59, SD = 0.50). Five schools were assigned to the Kalulu intervention group or a "business-as-usual" control group, and evaluated two times in 2022. **Results:** Kalulu group showed significant progress, reading an average of 4.9 more words per minute compared to control classes. Improvements were observed in phoneme verbal fluency and a memory task.

Conclusion: Given the strong scientific support for phonics and our encouraging results, we urge Brazil's educational leaders, teachers, and parents to advocate for increased phonics instruction as part of early literacy methods. Explicit phonics instruction, like the Kalulu method, could provide essential help for overcoming the country's ongoing literacy challenges.

---

[1] To preserve anonymity, we will only provide the website information once published.

In 2000, the National Reading Panel published a comprehensive meta-analysis to resolve debates between advocates of whole-language and phonics reading methods. (National Reading Panel, 2000). The scientific consensus was clear: early, explicit, and systematic phonics instruction is the most effective way to teach literacy to all students (Ehri et al., 2001). For any nation, providing high-quality reading instruction is essential for individual success and societal progress. To achieve this, many countries are increasingly prioritizing phonics as a core element of their national reading programs (Castles et al., 2018; CSEN, 2019; Rose, 2006). In Brazil, however, phonics instruction has not yet become standard practice. This is largely due to a historical resistance to traditional teaching methods, which arose amid the positive expansion of educational access and the growing concern for supporting underprivileged communities. Supporters of constructivists pedagogy urged teachers to move away from phonics, dismissing it as too mechanical, and instead promoted whole-language methods that emphasized learning to read through exposure to meaningful text (Ferreiro, 1999; Goodman, 1967). While constructivist methods were well-intentioned in their goal to expand education to previously underserved populations, they overlooked mounting evidence against their effectiveness and ultimately ignored the science of reading. Constructivist theories still dominate reading instruction in Brazil today. This disconnect between the science of reading and classroom practices has become a critical issue, as reflected in Brazil's low ranking of 53rd out of 57 countries in the 2021 Progress in International Reading Literacy Study (PIRLS) (Mullis et al., 2023).

Aiming to help improve this scenario, we present the implementation and results of a year-long reading intervention testing the use of the Kalulu phonics method with 1st-grade students in São Paulo. The Kalulu method is an open-source phonics method that can be downloaded at (absent to preserve anonymity).

Brazil's current literacy challenge: political historical context

To understand Brazil's current literacy challenges and the absence of phonics in classrooms, it is important to consider the historical context of reading instruction and related debates. By the end of the Brazilian Empire in the 19th century, schools were few and teachers predominantly used synthetic phonics methods from Europe, emphasizing the instruction of grapheme-phoneme correspondences (GPCs) and syllabic reading (Mortatti, 2019). In the early years of the Brazilian Republic, São Paulo introduced the first Brazilian Portuguese reading instruction book based on the North American analytical method, which emphasized starting with whole words and then analyzing their parts. This approach marked a shift from traditional phonics-based instruction that would start with teaching a grapheme-phoneme correspondence followed by syllable and word reading, towards students to read whole words, followed by the study of the associated letters. Public administration teachers promoted the analytical method through various publications, eventually making it mandatory in São Paulo schools. However, many educators criticized it as less effective and more time-consuming than synthetic phonics, leading to a conflict between the two methods until the late 1920s. This conflict contributed to the use of the term "alfabetização" (literacy), defined as the ability to "read and write" (Mortatti, 2019), and popularly used to describe skills obtained, or not, by students.

Unable to resolve the debate between phonics and analytical methods, São Paulo schools downplayed the importance of a single approach. During this period, the book *ABC Tests to Verify the Maturity Necessary for Learning to Read and Write* (1934) was published (Lima, 2019), introducing tests to assess reading readiness and to organize students into homogeneous classes based on their results. This shifted the focus from debating reading methods to the goal of providing customized literacy instruction tailored to each child's developmental stage, a practice that remained popular until the late 1970s (Mortatti, 2019).

In the 1980s, two major developments significantly shaped Brazil's education system. First, the nation emerged from a political dictatorship with a focus on improving social services, sparking a national emphasis on education. Child labor was abolished[2], and elementary schooling became mandatory, resulting in a surge of new students and a move away from tailored educational approaches, which was no longer feasible. Second, constructivist theories gained prominence, heavily influenced by Argentinian psychologist Emilia Ferreiro's research on the psychogenesis of language (Lima, 2019). Additionally, Brazilian educator Paulo Freire, known for his seminal work, *Pedagogy of the Oppressed*, advocated for education as a means of liberation. His approach to adult literacy emphasized a more contextualized method of teaching reading. Freire's contextualized teaching methods, along with Ferreiro's constructivist approach, deeply influenced Brazilian educators for children as well as for adults. This led to the widespread adoption of constructivist and whole-language instruction, which quickly became the dominant literacy approach in Brazil (dos Santos et al., 2018; Mortatti, 2019; Rangel et al., 2017).

Constructivism posits that learners actively construct their understanding through experiences and reflection, with teachers guiding this process by fostering exploration and problem-solving rather than directly imparting knowledge. Amid Brazil's political upheaval, constructivism was embraced as a philosophy aimed at challenging traditional educational practices to empower citizens (Mortatti, 2006). Numerous articles, academic theses, books, videos, and methodological suggestions were published to promote constructivist ideas and integrate them into the public education system. Over time, constructivism transformed reading instruction from a focus on specific skills to prioritizing the social function of written language.

---

[2] Child labor has been illegal in the country since the enactment of the 1988 constitution.

Paulo Freire's return to Brazil after exile and his role as Secretary of Education for São Paulo from 1989 to 1992 cemented his influence, leading to the widespread adoption of reading instruction methods based on constructivist principles in schools (Mortatti, 2006). In practical terms, concepts of right and wrong in terms of reading instruction disappeared from teachers' vocabulary, and traditional assessments of students' reading progress were phased out. The use of grade-appropriate reading materials, known as "readers," was also discontinued, as constructivist theory viewed them as presenting an artificial and unnatural universe of texts for reading (Azenha, 2006). Freire opposed the use of "child's readers," arguing they reinforced a divide between "those who know and those who know nothing" (Freire, 2014). He advocated for using "real texts", believing they allowed students to learn ideas and feelings about themselves and the world. Freire's views laid the foundation for Brazil's strong preference for contextualized, experience-based literacy, as promoted by constructivist educators (Feitosa, 2016).

The current context

The two major changes of the 1980s have resulted in a dual reality: a positive national commitment to educating all children, hindered by outdated methods that fail to deliver quality reading education. In 2001, a bill assigned states and municipalities the responsibility for executing school content, while federal bodies were tasked with supervising goals such as raising the overall educational level, improving education quality, and reducing social and regional inequalities (LAW N° 010172, 2001). Guidelines were set for education management, financing, professional development, and best practices. In 2003, a research group emphasized evidence-based literacy practices, leading to the introduction of the National Common Curricular Base (NCCB) in 2017, which provided a standardized educational framework. In 2019, the NCCB established a National Literacy Policy focused on scientific approaches, culminating in the launch of the National Reading Panel in 2021

(Brasil, 2024). During this period, the Basic Education Assessment System (SAEB) was introduced as a key component of Brazil's National Education Quality Index (IDEB). SAEB assesses students' performance in Portuguese reading, writing, and mathematics, which, along with school's dropout rates, contributes to the IDEB index, measured on a scale from 0 to 10 Brazilian schools (IDEB, 2024).

However, over the past 20 years, literacy rates have steadily declined, a trend worsened by the COVID-19 pandemic. In 2021, data revealed that 40.8% (see Figure 1) of children aged 6 to 7 were unable to read or write by the end of 2nd grade, with 51% from the lowest SES population (the poorest 25%) (IBGE, 2021). Although Figure 1 highlights the impact of COVID-19 on reading performance, the situation before 2020 was not much better. We contend that inappropriate reading methods were already contributing to these difficulties, and the pandemic-induced school closures only exacerbated the problem.

INSERT FIGURE 1

Ceará, a success story for reading progress

We argue that the poor reading outcomes in Brazil stem from outdated constructivist methods, not solely socioeconomic status (SES) issues. This is demonstrated by the case of Ceará, one of Brazil's poorest states, ranked 24th out of 27 in GDP per capita in 2021 (IBGE, 2023b). In 2005, Ceará's low SES was reflected in its public schools' IDEB scores for early grades, placing the state 18th out of 27 (QEdu, 2024). In 2007, Ceará launched the "Literacy Program at the Right Age," a government-led initiative that provided literacy education through teacher training, experience-sharing, monitoring, and resources, including synthetic phonics instruction (Gusmão & Ribeiro, 2011; Governo do Estado do Ceará, 2024). By 2021, Ceará's literacy scores had improved dramatically, elevating the state's IDEB ranking to 2nd place, despite its continued low SES.

The current project

The Ceará example demonstrated the positive impact that science-based reading programs can have on literacy, even in low SES regions. However, this success also involved broader changes, making it challenging to isolate the specific effect of phonics instruction. To address this, the current project was designed to specifically test the effects of phonics instruction. To our knowledge, only one prior study in Brazil has shown promising results in reading for 9-year-olds after training teachers used a phonics method, unfortunately this study only assessed a single school (Andrade et al., 2014).

To test a fully decodable phonics program without licensing fees or high costs, we utilized the adapted open-source Kalulu method to Brazilian Portuguese. Originally a tablet application, Kalulu is designed for easy translation into alphabetic languages, allowing global educators to provide phonics instruction. It has previously demonstrated improvements in phoneme awareness, fluency, and reading comprehension as an add-on to normal instruction (Potier Watkins et al., 2020; Potier Watkins & Dehaene, 2023). Kalulu is thus well-suited for this project's goal of introducing and testing phonics instruction with minimal disruption to teachers' existing practices. Numerous studies have demonstrated the benefits of educational games, including their effectiveness in reading instruction and in developing country contexts (Cancer et al., 2020; Gharibi et al., 2022; Görgen et al., 2020; Gori et al., 2013; Johann & Karbach, 2020; Juhani Lyytinen et al., 2021; Neville et al., 2009; Pasqualotto et al., 2022; Patel et al., 2018a; Puhakka, 2015; Roberts, 2021; van de Ven et al., 2017; van Gorp et al., 2017).

Kalulu was adapted for Brazilian Portuguese and context («hidden» [Manuscript submitted for publication]). Kalulu provides explicit instruction on grapheme-phoneme correspondences (GPCs), considering their frequency and consistency. Each grapheme-phoneme correspondence is instructed in a 3-part series based on evidence-based practices (Figure 2.A).

The GPC is introduced by an older child who pronounces the sound. Pictures of distinct articulatory gestures underlying each phoneme facilitate the development of early reading (Boyer & Ehri, 2011; Castiglioni-Spalten & Ehri, 2003). This is followed by an example of the letter sound in a common word with the accompanying picture (e.g., "/b/ in ball"). Finally, the child is asked to trace the grapheme. This provides an additional motor code to support memory for grapheme (Bara et al., 2004, 2016; Bara & Gentaz, 2011; Longcamp et al., 2005). After the lesson, the child masters the GPC through mini-games (Figure 2.B) that require increasingly fast responses. This approach improves grapheme-phoneme consolidation and accelerates the development of the brain's visual word form (Brem et al., 2010; Lassault et al., 2022; Ojanen et al., 2015; Patel et al., 2018b). Detailed explanations of the games are provided on the project website (not provided for anonymity).

INSERT FIGURE 2

The primary goal of this study was to test the effectiveness of the Kalulu phonics method in improving reading abilities in Brazilian public schools. While all students continued with standard whole-language instruction, those in the Kalulu intervention included phonics instruction three times a week for 15 weeks, totaling about 45 sessions during regular reading periods. Our main hypothesis was that students in the Kalulu group would show greater improvements in reading compared to the business-as-usual (standard school routine) control group. Additionally, we hypothesized that Kalulu would enhance phonological verbal fluency, as phonics training improves sensitivity to letter sounds, compared to semantic verbal fluency. A secondary hypothesis suggested that the use of games for learning would improve working memory. Unlike typical classroom instruction, Kalulu phonics offers a structured approach

with consistent GPC mapping, repetition, and recall, which stimulates working memory retention for verbal information (Demoulin & Kolinsky, 2016).

Methods

Participants

Our lab collaborated with the Secretary of Education in Santo André to recruit public schools for the project, requiring at least two schools to be low SES with low academic performance as indicated by IDEB scores. Although Santo André is a well-developed city in São Paulo DC, its public schools averaged an IDEB score of 6.1, below the target of 6.9. Five schools (representing 10 classes) volunteered: three schools were assigned to the 'Kalulu intervention' group (two mid-SES and one low-SES), and two were placed in the 'business-as-usual' control group (one mid-SES and one low-SES) (QEdu, 2024). Assignment was done by flip of a coin ensuring that at least one of the low-SES and mid-SES schools would be in each group, with the third mid-SES assigned to the Kalulu Intervention.

222 students were initially reported by teachers as participants. However, 8 students were excluded due to cognitive, motor, or genetic impairments that would hinder participation. Ultimately, 184 children (97 girls) aged 6 to 8 years (M = 6.59, SD = 0.50) completed both pre- and post-tests. Legal guardians provided informed consent, as approved by the Ethics and Research Committee of the Federal University of ABC (CAAE: 88208918.4.0000.5594).

Materials

Kalulu phonics game application

The game features a user-friendly interface (Figure 2), with Kalulu the hare guiding players through lessons organized into thematic gardens and minigames. Players progress through one garden at a time, completing 3 to 4 lessons (GPCs) per garden, starting with common and frequently occurring GPCs and advancing to more complex ones. Each GPCs is presented in a

short lesson then followed up by three mini-games. The player must score 80% correct to pass the mini-game. Difficulty is adjusted by wins and losses, aiming to help each player understand and complete the mini-games with minimal outside aid. You can learn more about the game and find links for its download in different languages (not provided for anonymity). The game application was developed using the Godot engine (https://godotengine.org), and the game code for the application used in the current project is available (not provided for anonymity).

Kalulu paper-based games

In addition to the digital game, a series of paper-based activities were developed (Figure 3), including card games, game boards, and homework assignments. These materials align with the game's phonics progression and mimic its playful, game-like aspects. The paper-based activities help children practice GP associations, letter recognition, syllable composition, decoding, and reading short sentences.

INSERT FIGURE 3

Procedure

In early 2022, all five schools participated in a pre-test, with classes divided into experimental and control groups. Teachers from the experimental groups attended two training sessions covering the game's functions, year-long protocol, session frequency, classroom structuring, and support activities. Each teacher also met with a researcher to organize their classroom layout, simulating the project's implementation. Teachers were responsible for informing students, organizing small groups, and managing game time. Researchers also talked to teachers about phonics. It is important to note that all teachers expressed their concern that phonics were boring and not appreciated as a method for reading instruction. During their

normal reading instruction time, all teachers used books provided by the selected publishers from the state government focusing on whole-word learning.

Classroom spaces varied across the three schools. Stations with different activities were set up, including one station with the tablets. Each class was divided into 3-4 groups of students. These groups rotated between stations with one group using tablets while others engaged in paper-based activities. Students using the tablets wore headsets to minimize noise and accessed the game with individual passwords. After 20 minutes, groups rotated activities. Teachers observed and assisted with support activities while minimally supervising the Kalulu group. A project team member ensured consistent implementation across classes through frequent visits to the classes during the Kalulu activity period, as organized by the teacher. The intervention lasted from March to mid-November 2022, with a break for July holidays, totaling around 45 sessions over 15 weeks. Classes used Kalulu materials 2-3 times per week, with each session including 20 minutes of tablet gameplay followed by 40 minutes of paper activities. Post-test assessments followed the intervention.

Measures

This intervention also included tests used for another project collecting observation data and norming tests for memory abilities in children. Below we report only those tests from the battery that were used to test the Kalulu intervention.

Predictor covariable

Rapid automatized naming (RAN). RAN was tested at the pre session to be used as a covariable in our analyses. RAN requires the subject to rapidly retrieve and say the names of visually presented, familiar items in a serial array (Denckla & Rudel, 1976; Norton & Wolf, 2012). RAN has been shown to be a unique predictor of reading fluency, in particular for children from low-SES backgrounds (Arnell et al., 2009; Caravolas et al., 2019; Clayton et

al., 2020; Georgiou et al., 2016; Ozernov-Palchik et al., 2017). This task is closely related to the automaticity and fluency required for efficient reading, as both involve quick retrieval of phonological information from memory. The RAN Colors and Objects were administered, but we chose to analyze only the responses to the common objects task (50 items: bed, cat, hand, pencil, and sun). This decision was made due to difficulties in identifying the brown color in the color test. The raw score, representing the time taken to name all items, was scaled and centered on 0 for use as a predictor variable.

Intervention assessment tasks.

1-minute word reading. The 1-minute reading task was chosen as our primary measure of student reading improvement. Children were presented 66 words (3, 4 and 5 letters) common for the Brazilian context and given one minute to read as many words as possible (Rodrigues et al., 2015; Salles & Parente, 2007). This task is considered a good measure of fluency primarily because words cannot be guessed through context like with text reading. All the words presented used regular GPCs. The raw score collected was the number of words read correctly in one minute.

Phonological verbal fluency (PVF). Children were given one minute to name as many words as possible that started with the letter 'P'. Names and places were not accepted. This task required that the child retrieve the sound of the letter 'P', /p/, and then name words that made this same first sound. The raw score was the total number of items named (Gonçalves et al., 2013).

Forward and backward digit repetition task. The child first completed the forward digit span task. In this task, the child was required to verbally recall a sequence of numbers in the same order it was presented (the child hears "5, 9, 7", and must repeat "5, 9, 7"). Following this, they completed the backward digit span task, which required the child to recall the sequence in reverse order (the child hears "5, 9, 7", and must repeat "7, 9, 5"). This standardized task

included two sequences of 2 to 5 digits in each measure (Figueiredo & Nascimento, 2007), for eight test items in each task (8 being the maximum score). Each child was given two trials with feedback on a 2-digit sequence. Once the test started, the child only advanced to the next sequence if at least one item was repeated correctly. We collected percentage correct separately for forward and backward. The mean percentage correct across the two tasks was then recorded and used in our analysis.

Control assessment tasks

Semantic verbal fluency (SVF). Children were given one minute to name as many animal names as possible. This tests children's vocabulary and ability to recall from a semantic category. The raw score was the total number of items named (Gonçalves et al., 2013). This task was used as a control for PVF. Our intervention focusses on teaching children grapheme-phonemes and decoding and encoding these in syllables in words, not on building vocabulary. In fact, all vocabulary used in the program were familiar to children. We would expect that all children would improve in this task, as compared to the PVF task, which should be facilitated by improve grapheme-phoneme knowledge.

Letter name knowledge. In Brazilian schools, learning letter sounds is rare, but children do learn their ABCs and letter name recognition. Children were shown all letters of the alphabet shuffled except 'k', 'y', and 'w' (due to their low frequency in Portuguese) and asked to name the letter. The raw score was the total number of correctly named letters, with a maximum possible score of 21. We expected that all children would improve equally in this test.

Statistical analysis

Several factors guided our choice of model to use to assess pre to post gains comparing students' outcomes depending on their group (Kalulu, or 'business-as-usual' control). We also want to consider the variability created by the different teachers in the project. Teachers in

Brazil are largely free to teach using the methods that they want, and these teachers were also located in schools with different levels SES as reported. Finally, children in Brazil enter school with highly variable rates of pre knowledge from their home environment and if they regularly attended kindergarten. For this reason, it was important to include pretest, age, and performance on the RAN task as a key predictor in reading fluency. Taking these factors into account, we opted to use hierarchical mixed-effects model. A hierarchical considers the nested structure of the data, where students are nested within classes. Mixed-effects includes both fixed effects (predictors that are consistent across all observations) and random effects (accounting for variability at the group level, in this case, school classes). We used the R Statistical Software (v 4.1.2; R Core Team 2021), package glmmTMB for this analysis. We specified the following hierarchical mixed-effect model:

$$\text{Posttest}_{ij} = \beta0 + \beta1 \cdot \text{Pretest}_{ij} + \beta2 \cdot \text{Treatment }_{ij} + \beta3 \cdot \text{RAN\_Pre}_{ij} + \text{Age\_Pre}_{ij}) + b0_j + \epsilon_{ij}$$

where:

i. Post_{ij} is the post-intervention score for student i in class j.

ii. $\beta0$ is the fixed intercept.

iii. $\beta1$ is the fixed effect coefficient for the pre-intervention score (Pretest_{ij})

iv. $\beta2$ is the fixed effect coefficient for the treatment group (Treatment_{ij})

v. $\beta3$ is the fixed effect coefficient for the RAN score at pretest (Ran_Pre_{ij})

vi. $\beta4$ is the fixed effect coefficient for scaled age at pretest (Age_Pre_{ij})

vii. $b0_j$ is the random intercept for class j, representing the deviation of the class j intercept from the overall intercept $\beta0$.

viii. ϵ_{ij} is the residual error term for student i in class j

During the posttest data collection, we collected a high incidence of zero scores in outcome variables. When over 5% of students scored 0 on a task at the posttest, we extended this

model to include a zero-inflated negative binomial component. This extension allowed us to account for both overdispersion and the excess zeros observed in the data. This combined model includes:

- The same fixed and random effects as the initial hierarchical mixed-effects model.

- A zero-inflated component (ziformula = ~1) to model the probability of excess zeros.

- The negative binomial distribution (family = nbinom2) to handle overdispersion in the response variable.

All scripts and anonymous data can be found on the project OSF account at «hidden for purposes of anonymity». Results are reported as significant at p < .05.

Results

We first examined the normality of the outcome variables, finding that several distributions were not normal, as to be expected where many students demonstrated persistent difficulties in completing tasks. The RAN pretest predictor variable was missing for one individual. Since all their other outcome scores were collected, we imputed their missing score using a regression model based on age and class. Participant demographics, pre- and post-test means, normality measures, and group differences are presented in Table 1. We used the Shapiro-Wilk test to assess the normality of outcome distributions for all participants on each task. If the distribution was normal, we used a t-test to test for group differences. If the distribution was not normal, we applied the Wilcoxon rank-sum test instead.

The Kalulu group had significantly better pretest scores than the control group in the 1-minute word reading (Kalulu: M = 2.55, SD = 9.95 vs Control: M = 0.36, SD = 1.85; W = 4246.5, p = 0.03) and letter knowledge tasks (Kalulu: M = 15.37, SD = 6.78 vs Control: M = 12.05, SD = 7.18; W = 4901, p = 0.001). These differences could be attributed to the larger number of mid-SES students in the Kalulu intervention group (2 classes versus 1 in the control group),

including 7 mid-SES children who read over 17 words on the 1-minute word reading task, making them outliers at 2 SD from the mean. We repeated the analyses excluding these outliers, but the significance of our results remained unchanged (see data and scripts in OSF for outcomes with and without these outliers). Therefore, we included all students in the reported results. We chose this approach because using a general mixed model and entering pretest levels as a fixed effect should control for pretest group differences when estimating intervention effects. Group means and intervention effects are illustrated in Figure 4.

One minute reading

Consistent with Brazil's reading crisis, 21% of students did not read a single word at the posttest (Kalulu = 47%, Control = 53%). A zero-inflated negative binomial model was thus used to analyze the data. We report a significant intercept for the zero-inflation component (β = -1.46, p < 0.001), meaning that there were more students reading 0 words than what would be expected by chance and, importantly, confirming that zero-inflation was essential in the model.

The pretest word reading score was a significant predictor of how many words students could read after the intervention, meaning that students who read more words before the intervention tended to improve more (β = 0.03, SE = 0.01, z = 2.9, p = 0.004). Similarly, students with faster RAN scores at the start also showed better improvement in reading (β = -0.01, SE = 0.004, z = -2.59, p = 0.01). However, students' age did not have a significant impact on their reading outcomes.

Participation in the Kalulu program significantly increased the number of words read compared to the control group (β = 0.36, SE = 0.150, z = 2.37, p = 0.02). To better understand the practical impact of this intervention, we can exponentiate the coefficient β from the log scale to the original count scale. Exponentiating 0.36 gives approximately 1.43, indication

that students in the Kalulu group are expected to read 1.43 times more words than those in the control group. This translates to a 43% increase in the expected number of words read. Given the control group's baseline average of 11.4 words, this 43% increase means that students in the Kalulu group are likely to read about 4.9 additional words in one minute compared to their peers in the control group. This result highlights the meaningful improvement in reading performance associated with the Kalulu intervention.

Phoneme verbal fluency

Given the lack of explicit phonics instruction in Brazilian schools, 8% of students scored 0 on this task (Kalulu = 71%, Control = 29%). The highly significant intercept in the zero-inflation model (β = -3.59, SE = 0.80, z = -4.51, p < 0.001) confirmed the need to account for excess zeros. The number of 'P' words retrieved at pretest significantly predicted the number named at posttest (β = 0.13, SE = 0.03, z = 4.57, p < 0.001), as did the RAN score (β = -0.01, SE = 0.003, z = -3.58, p < 0.001). Age was not a predictor of outcome.

Children in the Kalulu group were significantly more likely to associate the letter 'P' with its sound and retrieve words starting with the /p/ sound (β = 0.27, SE = 0.10, z = 2.68, p < 0.01). We can use this to understand the practical impact of this improvement by exponentiating the coefficient β. Exponentiating 0.27 gives approximately 1.31, meaning that children in the Kalulu group are expected to name 1.31 times more words starting with /p/ compared to those in the control group.

Digit span

All students were able to repeat at least one sequence., meaning the addition of the 0-inflation component to our model was not needed. The pre-intervention score significantly predicted the post-intervention score (β = 0.55, SE = 0.05, z = 10.34, p < 0.001). However, neither RAN or age at pretest had a significant impact on outcomes.

Children that had used the Kalulu method were significantly more likely to have higher post-intervention scores (β = 0.04, SE = 0.02, z = 2.43, p = 0.02). When we transform this β to the original scale, we get an exponentiate of 0.04, meaning a 4% boost in the expected post-intervention digit span outcome for the Kalulu group over the control group,

Semantic verbal fluency

All but one student was able to name at least one animal. The number of animals named at pretest significantly predicted the post responses, (β = 0.41, SE = 0.07, z = 5.88, p < 0.001), as did a child's RAN score (β = -0.04, SE = 0.01, z = -2.97, p < 0.01). Age was not a predictor of change. As predicted, all students equally improve in the number of animals names at posttest (β = 0.99, SE = 0.59, z = 1.68, p = 0.09).

Letter name knowledge

All students were able to name at least one letter. The number of letters the child could name at pretest significantly predicted letter naming at posttest (β = 0.28, SE = 0.003, z = 5.56, p < .001) as did pretest RAN scores (β = -0.03, SE = 0.01, z = -3.16, p < 0.01). Again, age was not a significant predictor.

We did not find a significant advantage in the ability to name letters provided by the Kalulu intervention (β = -0.45, SE = 0.48, z = -0.93, p = .35).

INSERT FIGURE 4

Discussion

Brazilian history has seen numerous shifts in teaching methods, often driven by political and social changes, culminating in a move toward constructivist approaches since the 1980s (Lima, 2019; Mortatti, 2006). Today, exploratory, discovery, and constructivist methods persist despite the overwhelming scientific consensus favoring phonics and explicit

instruction (Ehri et al., 2001; Klahr & Nigam, 2004; National Reading Panel, 2000). these methods overlook evidence that self-directed learning is more effective when learners already have some prior knowledge (Kirschner et al., 2006). This is particularly problematic for students in Brazilian public schools, who often lack foundational skills and would benefit most from explicit, systematic, and structured instruction (Fletcher et al., 2021).

The primary goal of the current study was to demonstrate the importance of phonics instruction in early reading education. This was challenging because the existing classroom methods could not be replaced, only supplemented. All students received regular instruction from their teachers, but in the Kalulu intervention, teachers also set aside 40-minute periods, two to three times a week, for a rotating reading workstation focused on phonics instruction. Our first hypothesis was that students participating in the program, providing structured game-based phonics instruction, would improve specifically in decoding fluency. Our results support this hypothesis: children in the Kalulu group read nearly five more words per minute than those in the control group, after controlling for pretest scores, processing speed and verbal retrieval, and classroom environment. As expected, however, children in the Kalulu group did not significantly outperform the control group in letter name knowledge, with 90% of experimental students and 88% of control students able to name more than 80% of the letters. This highlights that knowledge of letter names alone does not extend to reading proficiency if letter sounds are not also learned.

These results support previous research on the Kalulu app in France (Potier Watkins et al., 2020; Potier Watkins & Dehaene, 2023), but they are not isolated to this particular phonics method. Other studies in different countries have pointed to positive intervention outcome with phonics games for reading skills using tablet games: Belgium (Vanden Bempt et al., 2021), India (Patel et al., 2022), Netherlands (van Uittert et al., 2022), Indonesia (Debataraja et al., 2023), Germany (Berkling et al., 2015) and the United Kingdom (Ahmed et al., 2020).

Even in China, where the primary writing system is logographic, game-based teaching of English as a second language has been successful (Xin et al., 2023). Collectively, these results show that quality software with a clear reading goal can be beneficial in reading education. It's also worth noting that pen-and-paper games have proven to be effective in improving reading skills (Chan et al., 2023; Dessemontet et al., 2019; Ehri et al., 2001; Galuschka et al., 2020; Suggate, 2016).

The outcomes of this project are crucial for Brazil. Success or failure in acquiring reading skills directly impacts children's development. Many students who struggle with reading early on continue to face difficulties throughout their school years, with these challenges compounding over time (Carlson & Francis, 2002). A review of over 16,000 individuals (children, adolescents, and adults) found that poor readers are at a higher risk of developing anxiety and depression compared to typical readers at various stages of life (Francis et al., 2019). Therefore, quality reading instruction is essential to maintaining students' engagement and motivation. Since motivation significantly influences learning (Filgona et al., 2020) and emotions impact the consolidation of important memories (McClay et al., 2023), providing effective reading instruction is critical for literacy, but also keeping kids in school and providing them with the foundational skills to pursue educational goals.

Providing explicit phonics instruction is particularly important in Brazil, especially in low SES public schools, where students are often not tested or given assistance for learning disabilities. In such cases, quality phonics instruction becomes even more crucial. A meta-analysis of randomized controlled trials has shown that phonics instruction is the most effective method for improving reading and spelling performance not only for typical learners but is especially critical for students with reading disabilities, such as dyslexia (Galuschka et al., 2014).

As hypothesized, the Kalulu based intervention also improved students' results on the phonological verbal fluency task (PVF, naming as many words as possible that start with the letter 'p'), more than on the semantic task (SVF, naming animals). Kalulu's impact on phonological fluency may indicate an improved mental organization of phonological representations in these children. In this way, Kalulu affected not only reading but also broader phonological skills, aligning with other studies that have used game-based interventions (Amorim et al., 2020; Schmitt et al., 2018). Since initial reading acquisition heavily depends on phonological skills (see review in Melby-Lervåg et al., 2012), early phonological training can have a significant impact on both reading and broader phonological skills, particularly for disadvantaged children (Layes et al., 2022; Wolff & Gustafsson, 2022).

The secondary hypotheses tested in this study was related to children's working short term memory (VeSTM), assessed by the digits span task. Children in the Kalulu intervention showed better outcomes. These findings are important because VeSTM is not only related to reading comprehension and decoding (Nouwens et al., 2021) but also overall academic attainment (Alloway & Alloway, 2010). Moreover, some studies suggest that the relationship between reading and VeSTM is reciprocal (Miller-Cotto & Byrnes, 2020; Peng & Kievit, 2020). Considering the VeSTM outcomes alongside the background of the participating children, including their socioeconomic level and the pedagogic aspects of Brazilian public schools, all these factors may contribute to the persistent failure of literacy efforts in Brazil. The effect of socioeconomic level on VeSTM has been highlighted by a meta-analysis linking poverty to low working memory scores in developing countries (Nugroho et al., 2023). Additionally, children with low VeSTM scores, or smaller memory spans, seem to benefit more from interventions (for review see Diamond & Ling, 2020). Therefore, it's possible that children who played Kalulu improved their VeSTM due to the game's explicit training of strategies and composition, which required players to retain sounds (phonemes) and word

spellings in memory while strategically dismissing distractions, all within the time constraints of the game's animations. Specific training of working memory, combined with executive functions, has shown benefits in Brazilian public school students from low SES backgrounds, but not in private school students (Weissheimer et al., 2020). This is consistent with findings of low executive function skills in low SES communities, detectable even before formal education begins (Lipina et al., 2005). It should be studied future iterations of Kalulu if these benefits came from the game or specifically from learning how to read.

Limitations and Future Research Directions

Several limiting factors can be found in this work. First, the present study involved only five schools, with uneven numbers of students coming from low and mid SES backgrounds. In the future, a randomized control trial, with a larger and more diverse sample of schools would be important to truly capture the potential benefits of the program.

Another limitation fell with our ability to communicate and train teachers to the importance of phonics (even so, there was an effect of the intervention). Kalulu was integrated into the school routine with 2 to 3 sessions of 40 minutes per week, while the whole language method continued to dominate the rest of the school period. The teachers were largely unfamiliar with phonics methods and digital tools, which sometimes led to reduced involvement and interest in the project. When our team presented the positive results at the end of the program, we encountered skepticism about their validity.

Moving forward, it is essential to improve communication and provide comprehensive teacher training. To ensure successful implementation, it is crucial to involve school administrators and directors in supporting the shift to phonics instruction. This shift should be framed not as an admission of past mistakes, but as a progressive change to enhance student outcomes.

Many teachers still perceive phonics as an "old and boring" method. Overcoming this perception through effective communication and training is key to gaining acceptance. It's

vital that teachers feel excited and confident about the knowledge, skills, and tools they bring to the classroom, as their enthusiasm directly influences student engagement and success. These two factors combined also underpin an important criticism of the project, which is that at the end of the year, 16% of students in the Kalulu intervention were unable to read 1 word on our test. While this almost half of the number of students in the control (30% of students were unable to read 1word), we would have hoped that all students that used Kalulu might have started decoding the highly regular words presented on the test list. We believe that this could have been because both the computer and the card and board games were overly focused on the task of listening to a syllable or word and matching it to a written stimulus, a task that did not require students to read aloud. Reliable child voice recognition is not available, so Kalulu never requires that the child read aloud. All the games are based on hearing a syllable or word and matching it to its written form. This was similar in the group paper-based games, children heard words or saw pictures and then wrote them out or marked them on their bingo card. The game board required silent reading. These types of hearing and matching tasks were used because they were easy to coordinate in a group environment. Reading aloud exercises would have required that an adult take the time to listen and correct the child while they decode. This is often the job of the teacher. As our teachers did not have experience with asking children to decode words, this common and important task was not adequately practiced. It is our belief that while students using Kalulu may have learned a lot about the GPCs and spelling, we might have failed to boost reading aloud, or even the confidence to attempt reading aloud.

These last two criticism we hope to correct in future research. Today, Kalulu has been adapted to a full reading method with books for reading aloud and importantly, teacher training. Any future intervention would work to develop acceptability by teachers to use the complete

phonics method and, critically, provide more training for teachers to understand how they can support their students' developing reading skills.

Conclusion

The results following the use of the Kalulu phonics game showed a significant improvement in children's reading performance, consistent with outcomes observed in France and supported by the broader literature on phonics interventions. Given Brazil's unique socioeconomic and educational context, integrating Kalulu into schools as a literacy support tool could help bridge the gap between struggling readers and their peers. Further research and robust data are essential to persuade teachers and educational stakeholders of the critical need for targeted reading interventions. However, Kalulu alone is not sufficient; teachers play a pivotal role in promoting literacy effectively. Learning to read must be embedded within the curriculum, not treated as an add-on.

Acknowledgments

Funding:

Disclosure of interest

The authors report there are no competing interests to declare.

Data access statement

The data that support the findings of this study are openly available in Figshare at

https://figshare.com/s/9949f50e642b2b3f6e7a.

References

Ahmed, H., Wilson, A., 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 Education*, 5. https://www.frontiersin.org/articles/10.3389/feduc.2020.00132

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Alqahtani, S. S. (2020). Technology-Based Interventions for Children with Reading Difficulties: A Literature Review from 2010 to 2020. *Educational Technology Research and Development*, *68*(6), 3495–3525. https://doi.org/10.1007/s11423-020-09859-1

Amorim, A., Jeon, L., Abel, Y., Felisberto, E., Barbosa, L., & Dias, N. (2020). Using Escribo Play Video Games to Improve Phonological Awareness, Early Reading, and Writing in Preschool. *Educational Researcher*, *49*, 0013189X2090982. https://doi.org/10.3102/0013189X20909824

Andrade, E. M. de A., Mecca, T. P., Almeida, R. P. de, & Macedo, E. C. de. (2014). Eficácia de um programa de intervenção fônica para crianças com dificuldades de leitura e escrita. *Revista Psicopedagogia*, *31*(95), 119–129.

Arnell, K. M., Joanisse, M. F., Klein, R. M., Busseri, M. A., & Tannock, R. (2009). Decomposing the relation between Rapid Automatized Naming (RAN) and reading ability. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *63*(3), 173–184. https://doi.org/10.1037/a0015721

Azenha, M. da G. (2006). *Construtivismo. De Piaget A Emilia Ferreiro* (Vol. 8). Ática.

Bara, F., & Gentaz, E. (2011). Haptics in teaching handwriting: The role of perceptual and visuo-motor skills. *Human Movement Science*, *30*(4), 745–759. https://doi.org/10.1016/j.humov.2010.05.015

Bara, F., Gentaz, E., Colé, P., & Sprenger-Charolles, L. (2004). The visuo-haptic and haptic exploration of letters increases the kindergarten-children's reading acquisition. *Cognitive Development*, 433–449. https://doi.org/10.1016/j.cogdev.2004.05.003

Bara, F., Morin, M.-F., Alamargot, D., & Bosse, M.-L. (2016). Learning different allographs through handwriting: The impact on letter knowledge and reading acquisition. *Learning and Individual Differences*, *45*, 88–94. https://doi.org/10.1016/j.lindif.2015.11.020

Berkling, K., Pflaumer, N., & Lavalley, R. (2015). *German phonics game using speech synthesis—A longitudinal study about the effect on orthography skills*. Slate. https://www.semanticscholar.org/paper/German-phonics-game-using-speech-synthesis-a-study-Berkling-Pflaumer/6ad2633a8a57d7e6791b17a5278ccae60bb96f39

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*. https://www.tandfonline.com/doi/abs/10.1080/10888438.2010.520778

Brasil. (2024). *Ministério da Educação. Conheça a história da educação brasileira*. [Ministry of Education. Learn about the history of Brazilian education] http://portal.mec.gov.br/pet/33771-institucional/83591-conheca-a-evolucao-da-educacao-brasileira

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., Lyytinen, H., Brandeis, D., & Richardson, U. (2010). Brain sensitivity to print emerges when

children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences*, *107*(17), 7939–7944. https://doi.org/10.1073/pnas.0904402107

Cancer, A., Bonacina, S., Antonietti, A., Salandi, A., Molteni, M., & Lorusso, M. L. (2020). The Effectiveness of Interventions for Developmental Dyslexia: Rhythmic Reading Training Compared With Hemisphere-Specific Stimulation and Action Video Games. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01158

Caravolas, M., Lervåg, A., Mikulajová, M., Defior, S., Seidlová-Málková, G., & Hulme, C. (2019). A Cross-Linguistic, Longitudinal Study of the Foundations of Decoding and Reading Comprehension Ability. *Scientific Studies of Reading*, *23*(5), 386–402. https://doi.org/10.1080/10888438.2019.1580284

Carlson, C. D., & Francis, D. J. (2002). Increasing the Reading Achievement of At-Risk Children Through Direct Instruction: Evaluation of the Rodeo Institute for Teacher Excellence (RITE). *Journal of Education for Students Placed at Risk (JESPAR)*, *7*(2), 141–166. https://doi.org/10.1207/S15327671ESPR0702_3

Castiglioni-Spalten, M. L., & Ehri, L. C. (2003). Phonemic Awareness Instruction: Contribution of Articulatory Segmentation to Novice Beginners' Reading and Spelling. *Scientific Studies of Reading*, *7*(1), 25–52. https://doi.org/10.1207/S1532799XSSR0701_03

Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Chan, E. S. M., Shero, J. A., Hand, E. D., Cole, A. M., Gaye, F., Spiegel, J. A., & Kofler, M. J. (2023). Are Reading Interventions Effective for At-Risk Readers with ADHD? A Meta-Analysis. *Journal of Attention Disorders*, *27*(2), 182–200. https://doi.org/10.1177/10870547221130111

Clayton, F. J., West, G., Sears, C., Hulme, C., & Lervåg, A. (2020). A Longitudinal Study of Early Reading Development: Letter-Sound Knowledge, Phoneme Awareness and RAN, but Not Letter-Sound Integration, Predict Variations in Reading Development. *Scientific Studies of Reading*, *24*(2), 91–107. https://doi.org/10.1080/10888438.2019.1622546

CSEN. (2019). *ÉvalAide, un dispositif scientifique de prévention des difficultés en lecture et en mathématiques en CP et en CE1*. Canopé. https://www.reseau-canope.fr/fileadmin/user_upload/Projets/conference_role_experimentation_domaine_educatif/EvalAide_CSEN_Definitif_Mai2019.pdf

Debataraja, M. K., Heng, P. H., & Tiatri, S. (2023, February 20). *THE INFLUENCE OF PHONICS AND GAMES INTEGRATIVE LEARNING ON KINDERGARTEN CHILDREN'S READING READINESS | International Journal of Application on Social Science and Humanities*. https://journal.untar.ac.id/index.php/ijassh/article/view/25942

Demoulin, C., & Kolinsky, R. (2016). Does learning to read shape verbal working memory? *Psychonomic Bulletin & Review*, *23*(3), 703–722. https://doi.org/10.3758/s13423-015-0956-7

Denckla, M. B., & Rudel, R. G. (1976). Rapid "automatized" naming (R.A.N): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*(4), 471–479. https://doi.org/10.1016/0028-3932(76)90075-0

Dessemontet, R. S., Martinet, C., de Chambrier, A.-F., Martini-Willemin, B.-M., & Audrin, C. (2019). A meta-analysis on the effectiveness of phonics instruction for teaching decoding skills to students with intellectual disability. *Educational Research Review*, *26*, 52–70. https://doi.org/10.1016/j.edurev.2019.01.001

Diamond, A., & Ling, D. S. (2020). Review of the evidence on, and fundamental questions about, efforts to improve executive functions, including working memory. In *Cognitive and working memory training: Perspectives from psychology, neuroscience, and human development* (pp. 143–431). Oxford University Press. https://doi.org/10.1093/oso/9780199974467.003.0008

dos Santos, D. da S. T., Vidal, E. R. S., & Gutiérrez, A. J. C. (2018). ANALYSIS OF THE QUALITATIVE ASPECTS AND THE EFICIENCY OF LITERACY IN BRAZIL. *Pedagogia em Ação*, *10*(2). https://www.dropbox.com/s/uokshoyjd6y82nt/2022_scaff_jcl.pdf?dl=0

Ecalle, J., Kleinsz, N., & Magnan, A. (2013). Computer-assisted learning in young poor readers: The effect of grapho-syllabic training on the development of word reading and reading comprehension. *Computers in Human Behavior*, *29*(4), 1368–1376. https://doi.org/10.1016/j.chb.2013.01.041

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic Phonics Instruction Helps Students Learn to Read: Evidence from the National Reading Panel's Meta-Analysis. *Review of Educational Research*, *71*(3), 393–447. https://doi.org/10.3102/00346543071003393

Fälth, L., Gustafson, S., Tjus, T., Heimann, M., & Svensson, I. (2013). Computer-assisted Interventions Targeting Reading Skills of Children with Reading Disabilities – A Longitudinal Study. *Dyslexia*, *19*(1), 37–53. https://doi.org/10.1002/dys.1450

Feitosa, S. C. S. (2016). *Método Paulo Freire, interfaces e atualidade*. [Paulo Freire method, interfaces and current affairs]. https://acervo.paulofreire.org/handle/7891/4230

Ferreiro, E. (1999). *Psicogênese da Língua Escrita* (1ª edição). [Psychogenesis of Written Language]. Penso.

Figueiredo, V. L. M. de, & Nascimento, E. do. (2007). Desempenhos nas duas tarefas do subteste dígitos do WISC-III e do WAIS-III. *Psicologia: Teoria e Pesquisa*, *23*, 313–318. https://doi.org/10.1590/S0102-37722007000300010

Filgona, J., Sakiyo, J., Gwany, D., & Okoronka, A. (2020). Motivation in Learning. *Asian Journal of Education and Social Studies*, *10*, 16–37. https://doi.org/10.9734/AJESS/2020/v10i430273

Fletcher, J. M., Savage, R., & Vaughn, S. (2021). A commentary on and the role of phonics instruction in reading. *Educational Psychology Review*, *33*(3), 1249–1274. https://doi.org/10.1007/s10648-020-09580-8

Francis, D. A., Caruana, N., Hudson, J. L., & McArthur, G. M. (2019). The association between poor reading and internalising problems: A systematic review and meta-analysis. *Clinical Psychology Review*, *67*, 45–60. https://doi.org/10.1016/j.cpr.2018.09.002

Freire, P. (2014). *Pedagogia da esperança: Um reencontro com a pedagogia do oprimido*. [Pedagogy of Hope: A Reencounter with the Pedagogy of the Oppressed]. Editora Paz e Terra.

Galuschka, K., Görgen, R., Kalmar, J., Haberstroh, S., Schmalz, X., & Schulte-Körne, G. (2020). Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review. *Educational Psychologist*, *55*(1), 1–20. https://doi.org/10.1080/00461520.2019.1659794

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities: A Meta-Analysis of Randomized Controlled Trials. *PLOS ONE*, *9*(2), e89900. https://doi.org/10.1371/journal.pone.0089900

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2016). The anatomy of the RAN-reading relationship. *Reading and Writing*, *29*(9), 1793–1815. https://doi.org/10.1007/s11145-016-9653-9

Gharibi, H., Ahmadi, B., Mahmoodi, H., Karimian, E., & Gharibi, J. (2022). The Effectiveness of Attention Games on Reading Ability of Dyslexic Children. *Journal of Learning Disabilities*, *11*(3), 90–102. https://doi.org/10.22098/jld.2022.7263.1779

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, *29*(3), 207–219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Gonçalves, H. A., Mohr, R. M., Moraes, A. L., Siqueira, L. de S., Prando, M. L., & Fonseca, R. P. (2013). Componentes atencionais e de funções executivas em meninos com TDAH: Dados de uma bateria neuropsicológica flexível. *Jornal Brasileiro de Psiquiatria*, *62*, 13–21. https://doi.org/10.1590/S0047-20852013000100003

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Literacy Research and Instruction*. https://doi.org/10.1080/19388076709556976

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

Gori, S., Franceschini, S., Ruffino, M., Viola, S., Molteni, M., & Facoetti, A. (2013). Video Games Training Increases Reading Abilities in Children with Dyslexia. *Journal of Vision*, *13*(9), 282. https://doi.org/10.1167/13.9.282

Griffith, S. F., Hagan, M. B., Heymann, P., Heflin, B. H., & Bagner, D. M. (2020). Apps As Learning Tools: A Systematic Review. *Pediatrics*, *145*(1), e20191579. https://doi.org/10.1542/peds.2019-1579

Gusmão, J., & Ribeiro, V. (2011). Colaboração entre estado e municípios para a alfabetização de crianças na idade certa no Ceará. *Cadernos Cenpec | Nova Série*, *1*. https://doi.org/10.18676/cadernoscenpec.v1i1.37

Herodotou, C. (2018). Young children and tablets: A systematic review of effects on learning and development. *Journal of Computer Assisted Learning*, *34*(1), 1–9. https://doi.org/10.1111/jcal.12220

IBGE, I. B. de G. e E. (2021, February 1). *Impactos da pandemia na alfabetização de crianças*. https://todospelaeducacao.org.br/wordpress/wp-content/uploads/2022/02/digital-nota-tecnica-alfabetizacao-1.pdf

IDEB. (2024). *Índice de Desenvolvimento da Educação Básica (IDEB)*. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira | Inep. https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/ideb/indice-de-desenvolvimento-da-educacao-basica

Johann, V. E., & Karbach, J. (2020). Effects of game-based and standard executive control training on cognitive and academic abilities in elementary school children. *Developmental Science*, *23*(4), e12866. https://doi.org/10.1111/desc.12866

Juhani Lyytinen, H., Semrud-Clikeman, M., Li, H., Pugh, K., & Richardson, U. (2021). Supporting Acquisition of Spelling Skills in Different Orthographies Using an Empirically Validated Digital Learning Environment. *Frontiers in Psychology*, *12*, 566220. https://doi.org/10.3389/fpsyg.2021.566220

Kim, J., Gilbert, J., Yu, Q., & Gale, C. (2021). Measures Matter: A Meta-Analysis of the Effects of Educational Apps on Preschool to Grade 3 Children's Literacy and Math Skills. *AERA Open*, *7*(1). https://eric.ed.gov/?id=EJ1323865

Kirschner, P., Sweller, J., & Clark, R. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-

Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, *41*. https://doi.org/10.1207/s15326985ep4102_1

Klahr, D., & Nigam, M. (2004). The Equivalence of Learning Paths in Early Science Instruction: Effects of Direct Instruction and Discovery Learning. *Psychological Science*, *15*(10), 661–667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Lassault, J., Sprenger-Charolles, L., Albrand, J.-P., Alavoine, E., Richardson, U., Lyytinen, H., & Ziegler, J. C. (2022). Testing the Effects of GraphoGame Against a Computer-Assisted Math Intervention in Primary School. *Scientific Studies of Reading*, *26*(6), 449–468. https://doi.org/10.1080/10888438.2022.2052884

Layes, S., Guendouz, M., Lalonde, R., & Rebai, M. (2022). Combined Phonological Awareness and Print Knowledge Training Improves Reading Accuracy and Comprehension in Children with Reading Disabilities. *International Journal of Disability, Development and Education*, *69*(4), 1185–1199. https://doi.org/10.1080/1034912X.2020.1779914

LEI [LAW] N° 010172 (2001). http://portal.mec.gov.br/arquivos/pdf/L10172.pdf

Lima, F. R. (2019). ESTUDOS SOBRE ALFABETIZAÇÃO E LETRAMENTO NO BRASIL: Gêneses, desenvolvimentos e aplicações no ensino. *Cadernos Cajuína*, *4*(1), Article 1. https://doi.org/10.52641/cadcaj.v4i1.259

Lipina, S., Martelli, M., & Colombo, J. (2005). Performance on the A-not-B task of Argentinean infants from unsatisfied and satisfied basic needs homes. *Revista Interamericana de Psicología = Interamerican Journal of Psychology, ISSN 0034-9690, Vol. 39, N°. 1, 2005, Pags. 49-60*, *39*.

Longcamp, M., Zerbato-Poudou, M.-T., & Velay, J.-L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica*, *119*(1), 67–79. https://doi.org/10.1016/j.actpsy.2004.10.019

McClay, M., Sachs, M. E., & Clewett, D. (2023). Dynamic emotional states shape the episodic structure of memory. *Nature Communications*, *14*(1), 6533. https://doi.org/10.1038/s41467-023-42241-2

McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2020). 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

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

Miller-Cotto, D., & Byrnes, J. P. (2020). What's the best way to characterize the relationship between working memory and achievement?: An initial examination of competing theories. *Journal of Educational Psychology*, *112*(5), 1074–1084. https://doi.org/10.1037/edu0000395

Mortatti, M. D. R. (2006). *HISTÓRIA DOS MÉTODOS DE ALFABETIZAÇÃO NO BRASIL*. 16. https://nedeja.uff.br/wp-content/uploads/sites/223/2020/05/Histria-dos-Mtodos-de-Alfabetizao-no-Brasil.pdf

Mortatti, M. D. R. (2019). *Métodos de alfabetização no Brasil: Uma história concisa*. Editora UNESP. https://doi.org/10.7476/9788595463394

Mullis, I. V., von Davier, M., Foy, P., & Fishbein, B. (2023). *PIRLS 2021 International Results in Reading*. https://doi.org/10.6017/lse.tpisc.tr2103.kb5342

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*

(NIH Publication No. 00-4769). National Institute of Child Health & Development. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Neville, D. O., Shelton, B. E., & McInnis, B. (2009). Cybertext redux: Using digital game-based learning to teach L2 vocabulary, reading, and culture. *Computer Assisted Language Learning*, *22*(5), 409–424. https://doi.org/10.1080/09588220903345168

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, *63*, 427–452. https://doi.org/10.1146/annurev-psych-120710-100431

Nouwens, S., Groen, M. A., Kleemans, T., & Verhoeven, L. (2021). How executive functions contribute to reading comprehension. *The British Journal of Educational Psychology*, *91*(1), 169–192. https://doi.org/10.1111/bjep.12355

Nugroho, H. W., Salimo, H., Hartono, H., Hakim, Moh. A., & Probandari, A. (2023). Association between poverty and children's working memory abilities in developing countries: A systematic review and meta-analysis. *Frontiers in Nutrition*, *10*. https://www.frontiersin.org/articles/10.3389/fnut.2023.1067626

Ojanen, E., Ronimus, M., Ahonen, T., Chansa-Kabali, T., February, P., Jere-Folotiya, J., Kauppinen, K.-P., Ketonen, R., Ngorosho, D., Pitkänen, M., Puhakka, S., Sampa, F., Walubita, G., Yalukanda, C., Pugh, K., Richardson, U., Serpell, R., & Lyytinen, H. (2015). GraphoGame – a catalyst for multi-level promotion of literacy in diverse contexts. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00671

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D. E., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: Implications for early screening and theories of reading. *Developmental Science*, *20*(5), e12471. https://doi.org/10.1111/desc.12471

Pasqualotto, A., Altarelli, I., De Angeli, A., Menestrina, Z., Bavelier, D., & Venuti, P. (2022). Enhancing reading skills through a video game mixing action mechanics and cognitive training. *Nature Human Behaviour*, *6*(4), Article 4. https://doi.org/10.1038/s41562-021-01254-x

Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2018a). GraphoLearn India: The Effectiveness of a Computer-Assisted Reading Intervention in Supporting Struggling Readers of English. *Frontiers in Psychology*, *9*. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01045

Patel, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2018b). 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, P., Torppa, M., Aro, M., Richardson, U., & Lyytinen, H. (2022). 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

Peng, P., & Kievit, R. A. (2020). The Development of Academic Achievement and Cognitive Abilities: A Bidirectional Perspective. *Child Development Perspectives*, *14*(1), 15–20. https://doi.org/10.1111/cdep.12352

Potier Watkins, C., Caporal, J., Merville, C., Kouider, S., & Dehaene, S. (2020). Accelerating reading acquisition and boosting comprehension with a cognitive science-based tablet training. *Journal of Computers in Education*, *7*(2), 183–212. https://doi.org/10.1007/s40692-019-00152-6

Potier Watkins, C., & Dehaene, S. (2023). Can a Tablet Game That Boosts Kindergarten Phonics Advance 1st Grade Reading? *The Journal of Experimental Education*, *0*(0), 1–24. https://doi.org/10.1080/00220973.2023.2173129

Puhakka, C. S. A. (2015). Digital solutions for multilingual learning environments: The case of GraphoGame™ adaptations in Kenya. *Jyväskylä Studies in Education, Psychology and Social Research*, *525*. https://jyx.jyu.fi/handle/123456789/45762

QEdu. (2024). *QEdu—Data base.* https://qedu.org.br/

Rangel, F. de A., Souza, E. C. F. de, & Silva, A. C. de A. (2017). MÉTODOS TRADICIONAIS DE ALFABETIZAÇÃO NO BRASIL: PROCESSO SINTÉTICO E PROCESSO ANALÍTICO. *Revista Includere*, *3*(1), Article 1. https://periodicos.ufersa.edu.br/includere/article/view/7427

Roberts, T. A. (2021). Learning Letters: Evidence and Questions From a Science-of-Reading Perspective. *Reading Research Quarterly*, *56*(S1). https://doi.org/10.1002/rrq.394

Rodrigues, J. de C., Nobre, A. de P., Gauer, G., & Salles, J. F. de. (2015). Construção da tarefa de leitura de palavras e pseudopalavras (TLPP) e desempenho de leitores proficientes. *Temas em Psicologia*, *23*(2), 413–429. https://doi.org/10.9788/TP2015.2-13

Rose, J. (2006). *Independent review of the teaching of early reading: Final report.* DfES. http://dera.ioe.ac.uk/5551/3/5d970d28fc535dc54eb4dee9995bef36.pdf

Salles, J. F. de, & Parente, M. A. de M. P. (2007). Avaliação da leitura e escrita de palavras em crianças de 2ª série: Abordagem neuropsicológica cognitiva. *Psicologia: Reflexão e Crítica*, *20*, 220–228. https://doi.org/10.1590/S0102-79722007000200007

Schmitt, K. L., Hurwitz, L. B., Sheridan Duel, L., & Nichols Linebarger, D. L. (2018). Learning through play: The impact of web-based games on early literacy development. *Computers in Human Behavior*, *81*, 378–389. https://doi.org/10.1016/j.chb.2017.12.036

Suggate, S. P. (2016). A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions. *Journal of Learning Disabilities*, *49*(1), 77–96. https://doi.org/10.1177/0022219414528540

van de Ven, M., de Leeuw, L., van Weerdenburg, M., & Steenbeek-Planting, E. g. (2017). Early reading intervention by means of a multicomponent reading game. *Journal of Computer Assisted Learning*, *33*(4), 320–333. https://doi.org/10.1111/jcal.12181

van Gorp, K., Segers, E., & Verhoeven, L. (2017). Enhancing Decoding Efficiency in Poor Readers via a Word Identification Game. *Reading Research Quarterly*, *52*(1), 105–123. https://doi.org/10.1002/rrq.156

van Uittert, A., Verhoeven, L., & Segers, E. (2022). Responsiveness to a game-based intervention to enhance reading efficiency in first graders. *Journal of Computer Assisted Learning*, *38*(1), 178–191. https://doi.org/10.1111/jcal.12599

Vanden Bempt, F., Economou, M., Van Herck, S., Vanderauwera, J., Glatz, T., Vandermosten, M., Wouters, J., & Ghesquière, P. (2021). Digital Game-Based Phonics Instruction Promotes Print Knowledge in Pre-Readers at Cognitive Risk for Dyslexia. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.720548

Weissheimer, J., Fujii, R. C., & Souza, J. G. M. de. (2020). The effects of cognitive training on executive functions and reading in typically developing children with varied socioeconomic status in Brazil. *Ilha Do Desterro*, *72*, 85–100. https://doi.org/10.5007/2175-8026.2019v72n3p85

Wolff, U., & Gustafsson, J.-E. (2022). Early phonological training preceding kindergarten training: Effects on reading and spelling. *Reading and Writing*, *35*(8), 1865–1887. https://doi.org/10.1007/s11145-022-10261-x

Xin, G., Cheung, A., Abrami, P., & Wade, C. A. (2023). Examining the impact of ABRACADABRA (ABRA), a game-based online literacy program, on primary school

students in rural Hunan, China. *Educational Technology Research and Development*, *71*. https://doi.org/10.1007/s11423-023-10185-5

Table 1. Demographics and pre post outcomes for retained subjects in the Kalulu and the control group

	Kalulu	Control		Test for group differences
N (% girls)	115 (55%)	66 (50%)		
Age years:months	6:3 (0.3)	6:2(0.3)		t(133) = 2.14, p = .034
Pretest: Mean(SD)			**Shapiro-Wilk Test for Normality**	**Test for group differences** *
RAN in secs.	75.23 (26.43)	69.61 (12.42)	W = 0.77, p < 0.001	W = 4059, p = 0.44
1-min. word reading	2.55 (9.95)	0.36 (1.85)	W = 0.23, p < 0.001	W = 4246.5, p = 0.03
phoneme verbal fluency (PVF)	1.26 (1.6)	0.97 (1.26)	W = 0.73, p < 0.001	W = 4171.5, p = 0.24
Digit-span	0.38 (0.13)	0.39 (0.12)	W = 0.96, p < 0.001	W = 3598, p = 0.56
semantic verbal fluency (SVF)	8.82 (3.9)	9.45 (3.71)	W = 0.98, p < 0.02	W = 3482, p = 0.36
letter name knowledge	15.37 (6.78)	12.05 (7.18)	W = 0.84, p < 0.001	W = 4901, p = 0.001
Posttest: Mean(SD)				
1-min. word reading	20.3 (18.19)	11.39 (14.37)	W = 0.87, p < 0.001	W = 4992, p < 0.001
phoneme verbal fluency (PVF)	4.05 (2.8)	3.12 (2.34)	W = 0.93, p < 0.001	W = 4535.5, p = 0.03
Digit-span	0.50 (0.12)	0.47 (0.11)	W = 0.96, p < 0.001	W = 4173, p = 0.26
semantic verbal fluency (SVF)	11.11 (4.14)	10.58 (3.5)	W = 0.99, p = 0.48	t(154) = 0.93, p = 0.36
letter name knowledge	19.57 (3.89)	19.32 (3.34)	W = 0.47, p < 0.001	W = 4489, p = 0.01

* If data had a normal distribution, we compared group outcomes using a t-test. If data were not normal, we compared groups using a Wilcoxon rank sum test.
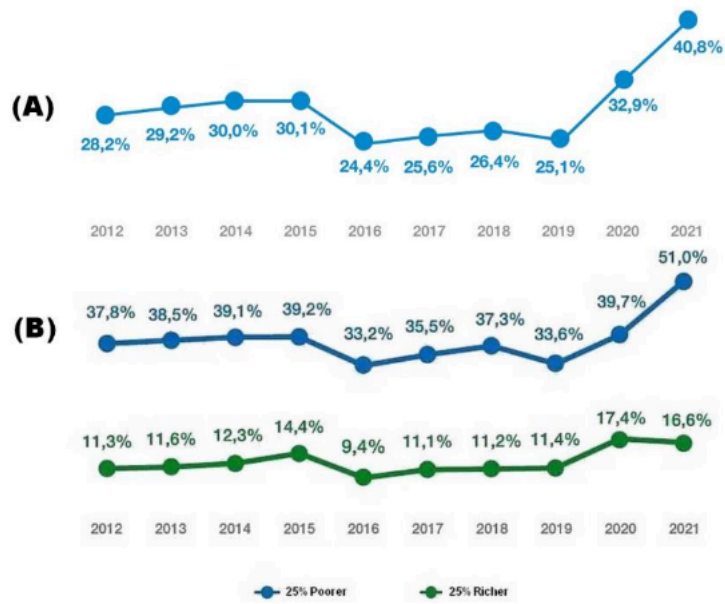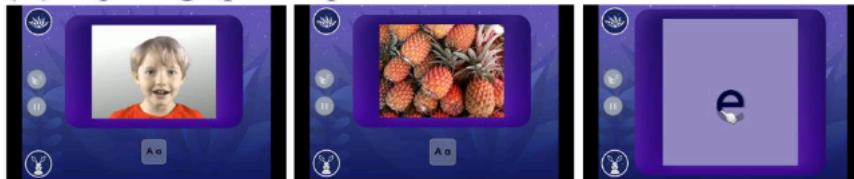
Figure 1

Figure 2

(A) Explicit grapheme-phoneme instruction

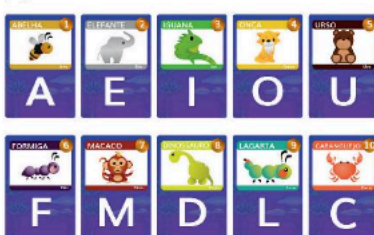(B) Mini-games to practice decoding and encoding

Figure 3

(A) NAMING CARDS

(C) GAME BOARD

(B) LETTER CARDS

(D) BINGO

Figure 4

a) Number of words read in 1 min
b) PVF, number of named 'P' words
c) Digitspan, % correct
d) SVF, number of named animals
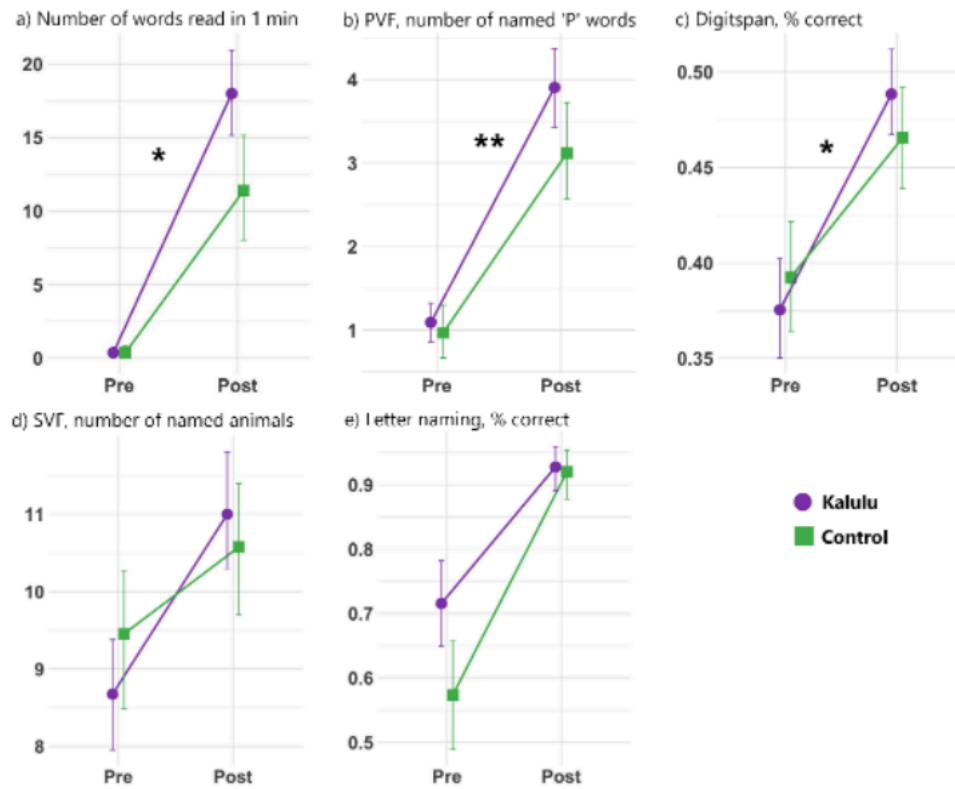e) Letter naming, % correct

Kalulu
Control

Figure Captions

Figure 1. Percentage of children aged 6 and 7 (tested at the end of 2nd grade, two years of formal reading education). These statistics are collected from national reading testes. (A) refers to the total percentage of children who cannot read and write in Brazil from 2012 to 2021. In (B), the blue line represents the illiteracy rates for children from the poorest quartile (per capita household income), and the green line represents the illiteracy rates for students from the wealthiest quartile (source IBGE, 2021). Illiteracy rates were high and stable since 2012 but saw a dramatic increase during the COVID pandemic.

Figure 2. Screenshots from the Kalulu game application. (A) In the Kalulu game application, each lesson starts with an explicit systematic phonics instruction in the following steps: a child demonstrates how to pronounce the lesson grapheme-phoneme correspondence, the phoneme is pronounced in a word with its associated picture and in the fashion "/a/ abacaxi", then the child traces the letter in upper- and lower- case. (B) Once the grapheme-phoneme lesson is done, the player is provided with three mini-games to master the lesson. Games require the player to engage in syllable decoding, spelling, and lexical decision. The game adjusts to the child's ability based on wins and errors made. Adjustments include changes in the number of areas from which stimuli can spawn, the number of distractors presented, and the speed required to click on the correct response before it disappears from the screen. All stimuli presented, including distractors, is decodable for the child according to the lessons completed.

Figure 3. Complementary cards and board games. Along with the tablet game app, students also played various phonics based card and board games including: (A) naming pictures then spelling the word using the cutout letter cards in panel, (B) using letter cards with images to remember letter sounds and using these cards to spell words; (C) matching words to their pictures on a board; (D) listening to a syllable or word and marking it on a Bingo card. This last game was played in small groups with the winner being the first to have marked five cells in a row. The most recent versions of these games can be downloaded at (this is kept hidden for anonymity).

Figure 4. Group Performance Improvements. Panels indicate performance in tests of (a) number of words read; (b) phoneme verbal fluency (ability to retrieve words starting with the letter 'P'); (c) forward and backward digit span repetition (d) semantic verbal fluency (number of animals named); (e) letter knowledge (number of letters named). All tests improved across time. Significance reported at $p < 0.05$*, $p < 0.01$**, $p < 0.001$***.