EVERFI®

An Efficacy Study of WORD Force

Early Literacy Digital Game-Based Learning Program Evaluation

Yang Bai, PhD, Jaumeiko Coleman, PhD, CCC-SLP, FNAP, Daniel Zapp, PhD

Executive Summary

The Research Team of EVERFI conducted an efficacy study of the WORD Force program in Spring 2024. Using an experimental design, the study recruited 320 students from families across the country and randomly assigned them to either the course group or the no-course group. Using the rigorous repeated measure mix model estimation method, the study shows the positive effects of WORD Force on several key aspects of early literacy skills on the students.

- **After the participants played at least six WORD Force games, their caregivers reported hearing a significantly wider variety of sentence structures their children's spoken language.**
- **The program effectively improved the participants' phonological awareness skills, as the caregivers reported a significant increase in their children's awareness of sound structures in spoken language after their children used WORD Force.**
- **The program significantly increased the participants' interest in words and reading-related activities.**
- **The study also evaluated the participants' early literacy skills using an external measurement tool, Star Early Literacy assessment, which is widely used in over 40 states and validated by numerous studies. The analysis showed that the participants who used WORD Force had a significantly larger improvement in their test scores than those who did not use the program.**

[1] **Acknowledgements:** This observational study was conducted by EVERFI, an educational technology company and the creators of WORD Force. The funding for this study was made possible by Truist.

Introduction

National data trends show that the majority of children in the United States continue to struggle to read proficiently (NAEP, 2022). This crisis was exacerbated by the COVID-19 pandemic and associated loss of instructional time, with our youngest students being most impacted (Amplify Education, 2022; Curriculum Associates, 2021; North Carolina State Board of Education and Department of Public Instruction, 2022).

Reading experts identified in the science of reading literature the following "Big 5" reading skills which children must learn to address the crisis and enable them to "read to learn": phonemic awareness, phonics, vocabulary, reading fluency, and reading comprehension (National Reading Panel & National Institute of Child Health and Human Development, 2000).

- **Phonemic Awareness** is the most advanced form of phonological awareness in which an awareness of and ability to manipulate the individual speech sounds in spoken words (i.e., phonemes) is present. For example, there are three phonemes in the word "team": /t/, /i/, /m/.
- **Phonics** is the ability to read and spell words using knowledge of the association between printed alphabet letters (and their combinations) and individual speech sounds (i.e., phonemes).
- **Vocabulary** learning is a component of semantics that reflects the degree of knowledge of a variety of words and the ways in which they can be used in language.
- **Reading Fluency** reflects the degree to which a person can read words with speed (i.e., automaticity), accuracy, and proper expression.
- **Reading Comprehension** occurs when an individual reads words fluently and concomitantly utilizes background knowledge, vocabulary, and reasoning skills to make meaning of text.

For young children nearing or just at the beginning of formal education, the pathway to achieving the Big 5 skills includes developing a strong foundation in early literacy skills like spoken language (including phonological awareness), concepts of print, early phonics, emergent spelling, and emergent writing (Armbruster, Lehr, Osborn, 2006; Armbruster, Lehr, Osborn, & Adler, 2006). A 2021 systematic review of early literacy instructional best practices by Herrera and colleagues revealed the following:

- Students best learn an early literacy skill like phonics through direct instruction.
- Teaching language 1:1 or in small groups appears to be most beneficial for students.
- In some cases, teaching multiple skills enhances learning of other skills; for example:
  - Print knowledge is enhanced when both print knowledge and phonological awareness are taught.
  - Decoding and early writing skills improve when print knowledge and phonics are taught.

National teacher shortages (Hammond et al., 2023; Learning Policy Institute, 2023), depressed national academic test scores (NAEP, 2022), and a continued focus on implementing science of reading-based instructional practices (Neuman et al., 2023) have compelled schools to consider additional ways to support student learning, such as ramping-up tutoring efforts (Hammond et al., 2023).

Researchers (Nickow et al., 2020) have noted the many benefits of tutoring, particularly high impact tutoring, which is provided by a well-trained educator in-person for three or more times per week, 30-minutes per session, in alignment with an evidence-based core curriculum (Institute of Education Sciences, 2022). However, some of the greatest barriers to in-person tutoring are difficulty finding trained tutors, lack of funding, tutors' lack of access to current performance data, students not showing up for tutoring, and time and space limitations (Institute of Education Sciences, 2022; National Student Support Accelerator, 2023).

Another mechanism for providing personalized instruction at scale is the use of digital game-based learning programs, which have been found to positively impact academic outcomes, including for younger students with and without learning disabilities (Behnamnia et al., 2023; Zolkipli et al., 2023). Digital game-based learning is often an immersive learning experience in which online games and associated elements, concepts, and mechanisms are integrated into school teaching and self-regulated learning activities (Pan et al., 2021).

Researchers have proposed that the look and feel of a game (aesthetics), storyline throughout the game (narrative), ways in which the game is navigated (game mechanics), and the disposition of the game player moderates a player's engagement and motivation to play the game. All of those factors collectively impact learning outcomes as indicated below:

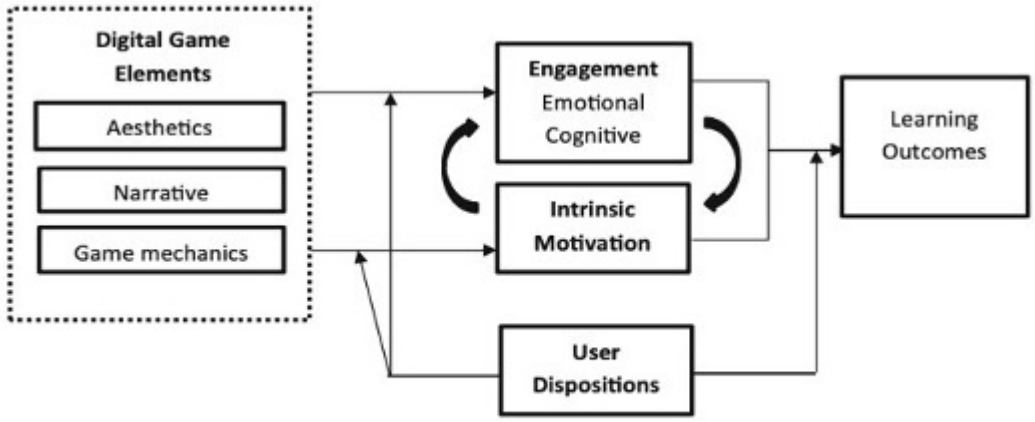


Fig. 1 A conceptual model linking game elements embedded in the instructional design with intrinsic motivation, user engagement and individual dispositions

Source: Alexiou, A. & Schippers, M.C. (2018). Digital game elements, user experience and learning: A conceptual framework. Education and Information Technologies, 23, 2545–2567. https://doi.org/10.1007/s10639-018-9730-6

As with any instructional program, the quality of the content and instructional methods is paramount (Liu et al., 2021). With regard to early literacy digital learning, programs that are engaging, motivating, and inclusive of practice on skills critical to literacy development (e.g., phonemic awareness, phonics, vocabulary, fluency, comprehension) have proven to enhance reading abilities (Putman, 2016).

While many studies of the impact of digital learning on early literacy have found positive impacts, they have not consistently provided information on recommended dosage (i.e., frequency, intensity, and duration of instruction) (Paul et al., 2023). There are mixed findings on the optimal dosage to achieve the highest outcomes from an early literacy digital game-based program (e.g., McTigue et al., 2019; Putman, 2016). A moderator analyses by Clark et al. (2016) of dosage features indicated that playing a digital game across multiples sessions (versus one session) was more positively impactful on learning outcomes than a specific duration.

Another consideration of playing digital games is recommendations on screen time, such as those from the American Academy of Pediatrics (AAP) (e.g., 1 hour a day for 2-5 year olds). International studies of actual screen time use tend to align with the AAP recommendation (Axelsson et al., 2022; Ribner & McHarg, 2021), which is important given research on developmental and behavioral outcomes associated with excessive screen time for children (McArthur et al. 2021).

Use of digital game-based learning programs by families could mitigate many of the barriers to in-person tutoring, while providing students with the academic support they need. Caregivers who have supported or observed their student playing digital game-based learning programs have noted engagement and positive outcomes (Ronimus et al., 2014; Xie et al., 2021). A host of research supports the importance and benefits of families actively supporting their children's literacy skill development (Terlitsky & Wilkins, 2015; Trinh et al., 2022), yet additional research is needed to learn more about factors that constitute a prime home literacy environment for effectively using game-based learning programs.

Given the breadth of support needed to address the reading acquisition crisis, an all-hands-on-deck approach is required, which includes using available digital learning programs that are proven to support reading acquisition. An observational study by Bai et al. (2022) of the digital game WORD Force examined its application in the classroom and relationship of those findings to home literacy environment practices. WORD Force, which has been reported by teachers to engage and motivate students to practice early literacy skills, addresses phonological awareness, phonics, vocabulary, and reading comprehension within a structured literacy framework (e.g., three layers of instructional support, adaptive pathing, pre- and post-assessments). The study results suggested the following:

- The larger the variety of games the students played, the greater their skill improvement was in the pre- to post-game assessments.
- Students with better in-game performance tended to have greater skill improvement.
- Compared to students who did all their engagement with the course in one day, those who consistently played a roughly equal number of games each day and spread out their usage across several days demonstrated greater skill improvement.
- Caregivers reported significant skill improvement for most students, with the greatest improvement reported being in phonological awareness and phonic skills among the pre-kindergarten students.
- The number of books available for children to read at home and the frequency with which caregivers taught their children how to read words were the most critical home literacy environment factors.
- In the classroom study, typical and struggling learners in pre-kindergarten, kindergarten, and first grade had greater improvement in the phonological awareness and phonics skills, while typical and struggling learners in the second grade had the greatest improvement in the spoken language skills. Compared to students who did all their engagement with the course in one day, those who consistently played a roughly equal number of games each day and spread out their usage across several days demonstrated greater skill improvement.
- All students had greater improvement if their teachers received training in teaching early literacy skills.
- Struggling learners, in particular, had greater skill improvement if their teachers used literacy blocks that included small group instruction, one-on-one instruction, independent work or peer work.

In an effort to explore other ways to expand literacy instruction support to more students, this study evaluated caregiver-supported implementation of WORD Force in the homes of pre-kindergarten through first grade students and its impact on their literacy skills. Factors evaluated in the Bai et al. (2022) observational study of WORD Force, such as home literacy environment practices, were also evaluated in this study. In addition, given the importance of understanding students' baseline literacy skills, caregivers proctored the pre- and post-administration of a nationally known online benchmark assessment of early literacy skills (i.e., Star Early Literacy from Renaissance Learning). Lastly, in consideration of dosage and screen time research and guidelines, students were asked to complete approximately seven lessons a week, which is roughly 70-minutes weekly. That dosage, which is well under recommended screen time limits, allowed students to play a variety of early literacy games in WORD Force in a manner found in previous research to have a positive impact on early literacy outcomes. The research questions for this study follow:

Reading experts identified in the science of reading literature the following "Big 5" reading skills which children must learn to address the crisis and enable them to "read to learn": phonemic awareness, phonics, vocabulary, reading fluency, and reading comprehension (National Reading Panel & National Institute of Child Health and Human Development, 2000).

1. **Are there statistically significant differences in learner's early literacy skills before and after engaging with WORD Force content?**

2. **Compared to a control group measured at the same time and in the same manner, can these differences be attributed to the WORD Force experience?**

Methodology

The Sample

A total of 357 children from 31 states were recruited to participate in this study. By the end of the study, 37 of them did not complete all the required activities and therefore were not included in the final analysis. The final sample included 320 children from 228 families. About two-thirds of the students were in kindergarten (36%) and pre-kindergarten (Pre-K) (33%), followed by –first graders (28%) and a small number of home-schooled children (3%). There were more male participants (60%) in this study. One-third (35%) of the participants were students of color, and 6% of the 320 children were English learners. The study was open to children with all levels of early literacy skills, and the final sample included 28 children who were delayed or had at least one type of disability (i.e., developmental, communication, learning). According to the caregivers' report, the vast majority of the participating children (85%) did not take any early literacy training courses before the study.

Study Design

To answer our research questions, researchers followed a cohort of families with children aged 4-7 years in Pre-K through first grade, using WORD Force while measuring caregivers' behaviors and the literacy environment in the home. Families who had not started using the product with their children were recruited through our social media networks and partner organizations. After recruitment, families were randomly assigned to either the course or comparison (i.e., no course) group of the study. Families in the course group were given 4 weeks to play at least six games in WORD Force. Both before and after taking the program, the children took an online, digitized assessment to evaluate their early literacy skills (Star Early Literacy Assessment). Families in the comparison group delayed taking WORD Force until data collection was finished and their students took the Star Early Literacy assessment during the same two occasions as the students in the treatment families. All caregivers took a survey about their family socioeconomic conditions, demographics and home literacy environment (Family Literacy Environment Survey) in the beginning of the study and were asked to evaluate their children's early literacy skills using a checklist (Early Literacy Evaluation Survey) we provided twice, at the start and end of the study. The final sample included 140 (44%) children assigned to the comparison group and 180 (56%) in the course group.

WORD Force Design

WORD Force consists of 15 games focusing on five different skill sets, as depicted below. Before starting each game, learners take a pre-game assessment of their related skill development. Each game includes five levels of play with each subsequent level increasing in difficulty and after the five levels are completed, their skills are again measured through a post-game assessment. Each learner's improvement in skills is reflected by the changes in assessment scores. Early literacy skills covered in WORD Force:

- **Game Set 1** - Phonological awareness (rhyme, onset-rime, phonemic awareness) and phonics (letter-sound association)
- **Game Set 2** - Phonological awareness (phonemic awareness) and spelling/word building
- **Game Set 3** - Spelling/word building and vocabulary
- **Game Set 4** - Spelling/word building and reading comprehension
- **Game Set 5** - Vocabulary and reading comprehension

Measurments

STAR Early Literacy Assessment

STAR Early Literacy (SEL) is a computer-adaptive assessment designed to measure the early literacy skills of young children from Pre-K to third grade, assessing their grasp on spoken language and literacy concepts and skills that directly influence their future ability to read. One of its advantages is the short amount of time it takes—only about ten minutes for administration. It consists of a 27-item adaptive assessment organized into three broad domains (Word Knowledge and Skills, Comprehension Strategies and Constructing Meaning, and Numbers and Operations) and 10 sub-domains (Alphabetic Principle, Concept of Word, Visual Discrimination, Phonemic Awareness, Phonics, Structural Analysis, Vocabulary Sentence-Level Comprehension, Paragraph-Level Comprehension, and Early Numeracy).

Children's performance on each test item influenced the difficulty of the next item, meaning that the assessment was adaptive by item across broad and sub-domains. We provide definitions of the sub-domains as measured by STAR in the supplemental materials. Based on these assessments, Early Literacy scaled scores and Early Literacy Grade Equivalent scores are computed. Scaled scores are not standardized by age and grade and should be interpreted similarly to raw scores. A scaled score is calculated based on the difficulty of questions and the number of correct responses and is useful for comparing student performance over time and across grades. STAR Early Literacy scaled scores range from 300 to 900. A scaled score of 530 indicates sufficient kindergarten readiness (Renaissance Learning, 2022).

Family Literacy Environment Survey and Early Literacy Skill Evaluation Survey

To assess the family literacy environment and caregivers' reading behaviors, researchers adapted a survey developed by Lombardino, Lieberman, and Coleman (2020). CaregiverCaregivers were asked to answer questions about the demographics of the children, their social-economic status, and the home literacy environment (e.g., frequency caregivers taught their children how to read words/letters and the number of books the children had at home to read) (Bradley et al., 2001; Gay et al., 2020; Hamilton et al., 2016; Puglisi et al., 2017; Sénéchal, & Young, 2008).

The caregivers were asked to evaluate their children's early literacy skills using an evaluation checklist we provided. The spoken language and literacy skills captured in the evaluation survey are aligned with the critical components of early literacy development as suggested by the National Reading Panel (2000) and include the following areas. See the appendix for a detailed explanation of each skill.

- *Phonological Awareness*
- *Alphabet and Phonics Knowledge*
- *Spoken Language*
  - *Speech Sound Production*
  - *Sentence Structure Complexity and Variety*
  - *Grammar*
  - *Vocabulary Learning*
  - *Storytelling and Retelling*
- *Interest in Words and Reading*

Statistical Method

This study used the repeated measure mixed model method (RMM) to evaluate the impact of WORD Force on the participants' early literacy skills. Compared to more conventional methods such as ANCOVA tests and repeated measure ANOVA, the RMM method has several advantages. First, it is better suited to settings where the treatment could have impacted different individuals in different ways. Second, in addition to the comparison of post-course outcomes, the RMM method makes it easy to show the trend differences between different groups, which was more suitable for the current study. Third, as a regression-type of analysis, the RMM estimation has less restriction on the types of covariates that can be added to the model (Detry & Ma, 2016). The analysis showed the differences between the course group and the comparison group in how their early literacy skills changed over the course of the study, after controlling for the effects of other influential factors such as the family literacy environment and students' and caregivers' demographics and socioeconomic status.

The Baseline Results

The participants' early literacy skills were examined before they started the course. In general, the caregivers' evaluation of their children's skills showed that most children had satisfactory spoken language skills, as over 80% of the children often showed good mastery of speech and sound production skills, age-appropriate grammar usage, variety and complexity of sentence structure, storytelling/retelling skills, and vocabulary learning skills. Most children (77%) were fairly interested in words and reading. Phonological awareness and phonics seem to be the areas where most children were struggling with. In terms of phonological awareness, only 17% of the children possessed all the 7 age-appropriate phonological awareness skills measured. Neither did the caregivers have a high rating for their children's phonics skills, as only 28% of the participants mastered all 6 essential phonic skills for children their age. To mitigate the subjective bias in the caregivers' evaluation, we also asked the children to take a standardized, digitized, and comprehensive early literacy skill test. The results showed on average, the students had a passable score of 858 out of 1100, and the vast majority of them (92%) were either at or above the national benchmark for children at similar ages. We conducted baseline equivalence analysis to ensure that the course group and the comparison group did not differ significantly in their starting levels of skill development. Based on the analysis, the effect sizes of the differences were either small (hedge's $g < 0.05$) or medium (hedge's $g < 0.15$). Therefore, our sample met the baseline equivalence requirement which is a prerequisite for valid comparisons of the two groups.

Besides the children's early literacy skills, the study also surveyed the home literacy environment of the participants' families. Based on the survey, although only 38% of the caregivers had a habit of reading, over 75% of them often read stories to their children. Over 80% of the caregivers reported to often teach their children words and letters. Close to 80% of the caregivers said there were over 40 books for their children to read at home, and about 80% of them said they often engaged in extended conversations with their children. There were 51 multi-lingual families where English was not the only language used at home, but the vast majority of the families (96%) used English as the primary language at home. Notably, 89% of the caregivers had a bachelor's or a higher degree, which might to a certain degree limit the generalizability of the study's findings. Those home literacy environment factors, though not the focus of the study, could profoundly affect the children's early literacy skill development and thus, were included in the analysis as control variables.

The Effects of WORD Force

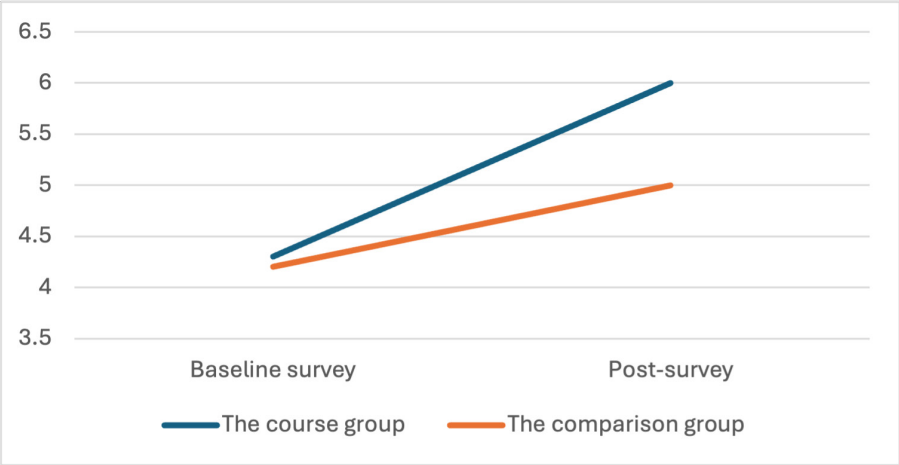
Using the repeated measure mixed model method (RMM), the impact of WORD Force on students' early literacy skills was analyzed. The analysis showed that WORD Force improved several critical early literacy skills for the participants based on caregivers' perspectives (i.e., Family Literacy Environment Survey) and results from a standardized assessment (i.e., Star Early Literacy). In the family literacy survey, caregivers indicated improved phonological awareness, sentence structure complexity and variety, and interest in words and reading. In addition, participants who used WORD Force had a significantly larger improvement in their Star Early Literacy scores than those who did not use the program.

Family Literacy Environment Survey Results

Phonological Awareness

Caregivers indicated that WORD Force had a significant positive impact on their children's phonological awareness. As shown in the chart below, controlling for the effects of other potential factors such as gender, race, and the family literacy environment, caregivers reported that participants' phonological awareness skills improved by 39% among those who played at least six WORD Force games, while the skills for those who did not use the program only improved by 16%.

CaregiverCaregivers' Evaluation of Children's Phonological Awareness (out of 7)

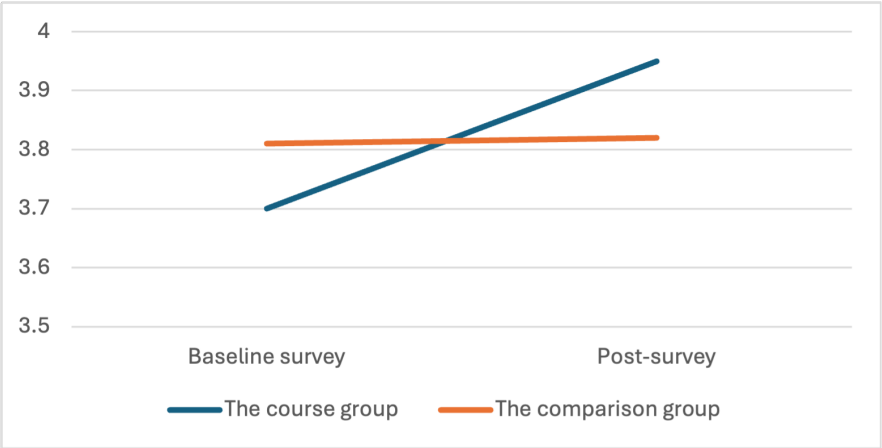


The effect of the course was further confirmed by the rigorous RMM analysis, which showed that the change trajectory of the course group was significantly different from that of the comparison group. In addition, the post-hoc trend analysis clearly indicated that although both groups had a positive change in their awareness of sound structures in spoken language, the improvement of the course group was significantly larger.

Sentence Structure Complexity and Variety

In the Family Literacy Environment Survey, we asked the caregivers to evaluate their children's ability to produce sentences that varied in length and complexity based on the variety of words used in different clause types . According to the RMM analysis and post-hoc trend comparison, WORD Force effectively improved the complexity and variety of sentences for the children. Specifically, as the chart below shows, controlling for the effects of other influential factors such as demographics and home literacy environment, the participants' levels of sentence structure complexity and variety perceived by their caregivers increased by 6% for those who played WORD Force games, while the caregivers' perception for those who did not use the program showed no changes.

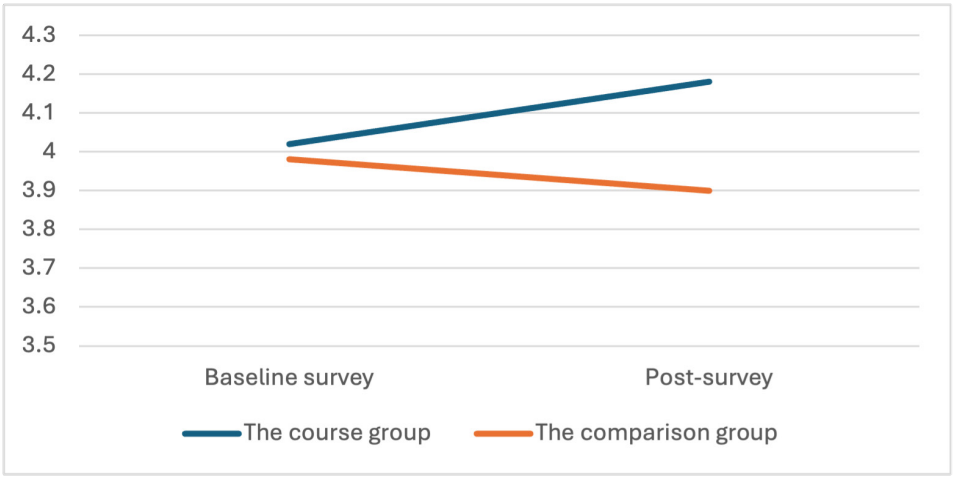
Caregivers' Evaluation of Children's Sentence Structure Complexity and Variety (out of 5)



Interest in Words and Reading

One objective of the WORD Force program is to stimulate students' interest in words and reading so that they will be motivated to continue to learn and improve their early literacy skills even after they use the program. The RMM analysis and post-hoc trend comparison clearly showed a significant difference between the course group and the comparison group in how the levels of interest in words and reading changed. Specifically, after playing at least six WORD Force games, the participants in the course group had a statistically significant increase (5%) in their interest in words and reading, as perceived by their caregivers, whereas those who did not use the program did not have any significant change in their interest levels. The levels of the children's interest in words and reading are shown in the chart below.

Caregivers' Evaluation of Children's Interest in Words and Reading (out of 5)



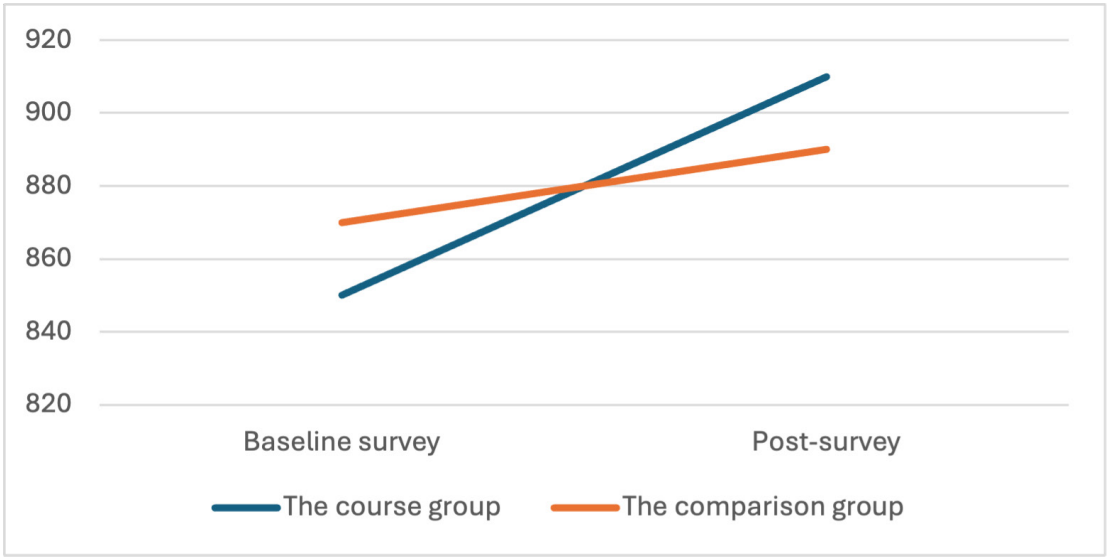
Although our analysis showed that WORD Force had a significant positive impact on several critical early literacy skills for those who used the program, there were a few skills that the program did not effectively improve, including the alphabet and phonics skills, grammar usage, speech and sound production and storytelling/retelling skills. The possible explanations for the lack of significant impact of the course on those skills are provided in the Conclusion and Discussion section.

Star Early Literacy Score

To provide a standardized comparison of early literacy skills, participants were administered the Star Early Literacy assessment, a formal test of early literacy skills validated and widely used in schools in over 40 states. The change in Star Early Literacy scores provided a more objective depiction of WORD Force's effect on the children's early literacy skills.

As the chart below shows, after controlling the effects of other influential factors, the average assessment score for the children in the course group increased by 7% from 850 to 910, while those who did not take the course only increased by 2%. Although the scores for both groups increased, the RMM analysis and the post-hoc comparison showed that the increase among the course group was significantly larger.

Star Early Literacy Score (out of 1100)



Conclusion and Discussion

Built on the observation study the EVERFI research team conducted in 2022, the efficacy study further tested the effects of WORD Force in improving the early literacy skills for young learners. Using the rigorous RMM analysis, the randomized controlled experimental study showed that the WORD Force program had a significant positive impact on several fundamental early literacy skills and provided compelling evidence that the improvement was caused by the course rather than natural skill acquisition over time or the impact of other influential factors such as the home literacy environment and/or caregivers' reading habits.

Based on survey responses from participating children's caregivers, our analysis showed that the children who played at least six WORD Force games had a significantly larger improvement in their phonological awareness than those who did not use the program. For this study, most participants played the first six games, which focus on phonological awareness and phonics skills. Therefore, a seemingly unexpected finding is that the games led to a significant improvement in the children's ability to use a variety of words in different sentence structures in their spoken language. However, the games played in the study included practice of advanced phonological awareness (i.e., phonemic awareness) skills and early conventional spelling patterns (e.g., cat; CVC), which have been found to positively impact spoken language skills when they are addressed in concert (e.g., Moxam, 2020).

The concomitant focus on phonemic awareness and phonics in the form of spelling also likely underlies the improvement in children's interest in words and reading. The study revealed that 77% of the participants already showed quite strong interest in words and reading prior to entering the study, so it was not anticipated that the program would make the participants more interested in words and reading. Yet, a significant improvement (5%) in interest in words and reading occurred. Researchers have noted that improved reading ability is associated with increased motivation to read (Wang et al., 2024). The spelling tasks in the games played for the study contain spelling patterns common in decodable books at a reading level appropriate for the age- and grade-level of study participants.

Besides the caregivers' evaluation of their own children's skills, the study also asked the children to take a standardized test, the Star Early Literacy assessment. The analysis showed that compared to the children who did not use WORD Force, those who played at least six games in the program had a significantly larger increase in their test scores. Both subjective (caregivers' observations) and objective (Star Early Literacy) measures of early literacy skill ability post-WORD Force use showed improvement in those skills. Therefore, the improved outcomes were unlikely to merely reflect the perceptions of caregivers about their children's skills but indicated real skill acquisition. Moreover, the Star Early Literacy assessment complemented the caregiver's survey, as the assessment covers a much wider range of early literacy skills.

The finding that the children who used WORD Force had a much larger increase in their Star Early Literacy assessment than noted by caregivers on the surveys is not surprising. The computer adaptive feature in the assessment allows for an ability-centered and nuanced, yet comprehensive, evaluation of a variety of academic skills (e.g., spoken language, literacy, numeracy). The caregiver surveys provide a general representation of perceived progress in broader areas of spoken language and literacy in a natural environment after completing WORD Force. Despite these discrepancies, both the assessment and caregiver survey results showed a positive impact from WORD Force.

There are a few skill areas which, based on the caregivers' perceptions, improved slightly but there was no significant difference based on completion of WORD Force. An exception was speech sound production, for which there was no notable change for either group. This makes sense given that direct instruction uniquely tailored to address specific sounds that are difficult for a child to produce is the gold standard for improving speech sound production (Farquharson & Tambyraja, 2022).

It is also worth mentioning that many students were reported to already have a good mastery of some of the skills even before the study. For example, as reported by their caregivers in the baseline survey, 86% of the children already mastered age-appropriate knowledge of a variety of words and how they are used in spoken language (vocabulary learning), 85% of them already showed age-appropriate skills to use of all parts of language to convey thoughts in an ordered and clear fashion (storytelling and retelling), and 86% of the children often showed age-appropriate grammatical skills in their speech. The lack of room for further improvement might explain why the course did not have a significant effect on those skills. Another possible explanation for the lack of significant impact of the program is that the children in the study were only asked to complete the first and second game sets. Had they completed later game sets, they would have practiced games that addressed vocabulary, grammar, and storytelling, and perhaps improvements in those skills would have been reported by caregivers of children in the control group.

Notably, all games in WORD Force feature fully voiced, animated characters who give users instruction and feedback. The cheerful sound of the characters may be a key factor in keeping students engaged in learning foundational reading skills. Therefore, as previous research (Behnamnia et al., 2023; Zolkipli et al., 2023) has shown, game elements (e.g., aesthetics, game mechanics) may have motivated and facilitated learning.

Another important factor is that while most students completed all six games across the two game sets studied, not all students completed all five levels in each game. In some cases that may have been due to the content in the later levels of a game being geared to an older child at a more advanced developmental level; in other cases, it could have been because the child needed more time to complete all the levels. Even still, students in the control group made significant progress on certain skills as indicated above during the 4-week study duration. Other researchers (e.g., Linstad et al., 2017) have found that study duration is an important component in facilitating learning, particularly for academics and language development. In light of research of this nature, the impact of WORD Force in the small amount of time it was used is even more profound.

Future research could extend the home study to include completion of the last three game sets in WORD Force by advanced students and those in higher grade levels. Testing students at the beginning of the school year, during the summer, and for longer periods of time could shed light on any time-specific instructional considerations for students at different developmental levels. In addition, disaggregated Star Early Literacy assessment data by spoken language and early literacy subskills would indicate which games in isolation and combination are most critical for students in different grade- and developmental-levels. This efficacy study of the impact of WORD Force when used in a home environment coupled with the findings from the observational study of WORD Force use in schools illuminates its positive role in early literacy skill acquisition.

References

Amplify Education. (2022). Research brief: Amid academic recovery in classrooms nationwide, risks remain for youngest students with least instructional time during critical early years. https://amplify.com/wp-content/uploads/2022/02/mCLASS_MOY-Results_February-2022-Report.pdf/

Armbruster, B. B., Lehr, F., & Osborn, J. (2006). Put reading first: The research building blocks for teaching children to read: Kindergarten through grade 3. National Institute for Literacy.

Armbruster, B. B., Lehr, F., Osborn, J., & Adler, C. R. (2006). A child becomes a reader: Proven ideas from research for parents-Kindergarten through grade 3 (3rd ed.). National Institute for Literacy.

Axelsson, E.L., Purcell, K., Asis, A., Paech, G., Metse, A., Murphy, D., & Robson, A. (2022). Preschoolers' engagement with screen content and associations with sleep and cognitive development. Acta Psychologica, 230. https://doi.org/10.1016/j.actpsy.2022.103762

Bai, Y., Coleman, J.J., & Zapp, D. (2022). An observational study of WORD Force: Early literacy digital game-based learning program evaluation. https://everfi.com/wp-content/uploads/2023/04/word-force-observational-study.pdf

Behnamnia, N., Kamsin, A., Ismail, M.A.B., Hayati, S.A. (2023). A review of using digital game-based learning for preschoolers. Journal of Computers in Education, 10(4), 603–636 https://doi.org/10.1007/s40692-022-00240-0

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), https://doi.org/10.3102/0034654315582065

Curriculum Associates. (2021). Understanding student learning: Insights from Fall 2021. https://www.curriculumassociates.com/-/media/_mainsite/_files/i-ready/iready-understanding-student-learning-paper-fall-results-2021.pdf

Farquharson, K., & Tambyraja, S. (2022). Introduction: Innovations in treatment for children with speech sound disorders. Language Speech and Hearing Services in Schools, 53, 627-631. https://doi.org/10.1044/2022_LSHSS-22-00065

Institute of Education Sciences. (2022). Tutoring. https://ies.ed.gov/schoolsurvey/spp/

Learning Policy Institute. (2023). The state of the teacher workforce: A state-by-state analysis of the factors influencing teacher shortages, supply, demand, and equity [Interactive map]. https://learningpolicyinstitute.org/product/state-of-teacher-workforce-interactive

Linstead, E., Dixon, D.R., Hong, E., Burns, C.O., French, R., Novack, M.N., & Granpeesheh, D. (2017). An evaluation of the effects of intensity and duration on outcomes across treatment domains for children with autism spectrum disorder. Translational Psychiatry, 7, e1234. https://doi.org/10.1038/tp.2017.207

Liu, Y.C., Wang, W.-T., & Lee, T.-L. (2021). An integrated view of information feedback, game quality, and autonomous motivation for evaluating game-based learning effectiveness. Journal of Educational Computing Research, 59(1) 3–40.

McArthur, B.A., Tough, S. & Madigan, S. (2021). Screen time and developmental and behavioral outcomes for preschool children. Pediatric Research. https://doi.org/10.1038/s41390-021-01572-w

McTigue, E., Solheim, O.J., Zimmer, W., & 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. https://doi.org/10.1002/rrq.256

National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.). (2000). Report of the National Reading Panel: Teaching children to read : an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health and Human Development.

National Student Support Accelerator. (2023). Challenges and solutions to implementing tutoring at scale. https://studentsupportaccelerator.com/sites/default/files/Challenges%20Solutions%20Implementing%20Turoing%20at%20Scale.pdf

North Carolina State Board of Education & Department of Public Instruction. (2022). Report to the North Carolina General Assembly: An impact analysis of student learning during the COVID-19 pandemic. https://content.govdelivery.com/attachments/NCSBE/2022/03/02/file_attachments/2091616/JLEOC%20Report%20 HB196.%20Impact%20on%20Lost%20Instructional%20Time%20for%20SBE%20March.pdf

Neuman, S.B., Quintero, E., & Reist, K. (2023). Reading reform across America: A survey of state legislation. Albert Shanker Institute. https://www.shankerinstitute.org/sites/default/files/2023-07/ReadingReform%20ShankerInstitute%20FullReport%20072723.pdf

Nickow, A., Oreopoulos,P., & Quan, V. (2023). The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence. National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w27476/w27476.pdf

Paul, C.D., Hansen, S.G., Marelle, C., & Wright, M. (2023). Incorporating technology into instruction in early childhood classrooms: A systematic review. Advances in Neurodevelopmental Disorders, 7, 380-391. https://doi.org/10.1007/s41252-023-00316-7

Putman, R. (2016). Technology versus teachers in the early literacy classroom: An investigation of the effectiveness of the Istation integrated learning system. Educational Technology Research and Development. https://doi.org/10.1007/s11423-01609499-5

Ribner, A.D., & McHarg, G. (2021). Screens across the pond: Findings from longitudinal screen time research in the U.S. and U.K. Infant and Behavior and Development, 63. https://doi.org/10.1016/j.infbeh.2021.101551

Ronimus, M., Kujala, J., Tolvanen, A., & Lyytinen, H. (2014). Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. Computers & Education, 71(3), 237-246. https://doi.org/10.1016/j.compedu.2013.10.008

Terlitsky, A.B., & Wilkins, J. (2015). Characteristics of family literacy programmes that improve child literacy, behaviour and parenting skills. International Journal of Pedagogies and Learning. https://doi.org/ 10.1080/22040552.2015.1113846

Trinh, N.T.M., Anh, N.T.T., & Hao, T.T.M. (2022). Integrating digital technologies into the early literacy learning of young children. Vietnam Journal of Educational Sciences, 18(2). https://doi.org/10.15625/2615-8957/22210204

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. National Assessment of Educational Progress (NAEP). (2022). Reading assessment.

Wang, T., Xu, H., Li, C., Zhang, F., & Wang, J. (2024). Dynamic insights into research trends and trajectories in early reading: An analytical exploration via dynamic topic modeling. Frontiers in Psychology, 15. https://doi.org/10.3389/fpsyg.2024.1326494

Xie, J., Wang, M., & Hooshyar, D. (2021). Student, parent, and teacher perceptions towards digital educational games: How they differ and influence each other. Knowledge Management & E-Learning, 13(2). https://files.eric.ed.gov/fulltext/EJ1314846.pdf

Zolkipli, N.Z., Rahmatullah, B., Samuri, S.M., Árva, V., & Pranoto, Y.K.S. (2023). 'Leave no one behind': A systematic literature review on game-based learning courseware for preschool children with learning disabilities. Southeast Asia Early Childhood Journal, 12(1), 79-97. Retrieved from: https://files.eric.ed.gov/fulltext/EJ1389218.pdf

Appendix

Operational Definitions[5]

Phonological Awareness is an umbrella term that refers to the awareness of the sound structure in spoken words at the word, syllable, rhyme, and phoneme levels (e.g., onset-rime awareness, phonemic awareness).

- **Onset-Rime Awareness** is a phonological awareness component in which there is an awareness of the onset (initial phoneme in a syllable) and the rime (the vowel phoneme and any subsequent consonant sounds) in a syllable. For example, the onset for the one syllable word "team" is /t/ and the rime is /im/.
- **Phonemic Awareness** is the most advanced form of phonological awareness in which an awareness and ability to manipulate the individual speech sounds in spoken words (i.e., **phonemes**) is present. For example, there are three phonemes in the word "team": /t/, /i/, /m/.

Phonics is the ability to read and spell words due to the knowledge that printed alphabet letters (and their combinations) represent individual speech sounds (i.e., **phonemes**).

Spoken Language is a two component skill: the ability to understand information communicated verbally and the ability to clearly and accurately express thoughts verbally to others. A few spoken language subskills captured in the family survey follow:

- **Speech Sound Production** is a component of *articulation* and *phonology together* which is the ability to form and verbally produce the individual speech sounds that make up spoken words in a language.
- **Sentence Structure Complexity and Variety** is a component of *syntax* and pertains to the ability to produce sentences that vary in length and complexity (e.g., number and type of clauses, use of advanced vocabulary).
- **Grammar i**s a component of *morphology* and *syntax* and signals awareness of the ways in which affixes (i.e., prefixes like "un" on "unlike" and suffixes like "er" on "runner") can change the tense and meaning of individual words and words in the context of sentences.
- **Vocabulary Learning** is a component of *semantics* and it reflects the degree of knowledge of a variety of words and the ways in which they can be used in language.
- **Storytelling and Retelling** allows the use of all parts of language (i.e., phonology, semantics, morphology, syntax, and pragmatics) to convey thoughts in an ordered and clear fashion.