International Journal of Research Studies in Education
2026 Volume 15 Number 12, 227-233

The A.I. KID study: Evaluating artificial intelligence interactivity and vocabulary acquisition in pre-kindergarten learners

Juan, Leandel U. ✉
Leyte Normal University, Philippines (2100962@lnu.edu.ph)

ISSN: 2243-7703
Online ISSN: 2243-7711

Received: 1 May 2026
Revised: 15 May 2026
Accepted: 16 May 2026
OPEN ACCESS
Available Online: 17 May 2026
DOI: 10.5861/ijrse.2026.26183

Abstract

This study evaluated the effectiveness of AI-assisted interactive applications on vocabulary acquisition among Pre-K children in Tacloban City, Philippines, using a quasi-experimental two-group pretest–posttest design. Vocabulary depth was measured using the Vocabulary Knowledge Scale (VKS; Wesche & Paribakht, 1996) among (n = 15) children in a morning class (Experimental Group) receiving ten (10) sessions of AI through the Buddy.ai, Google Read Along, and Readability apps, and (n = 15) children in an afternoon class (Control Group) receiving standard Pre-K instruction over the same two weeks. Both groups had Pre-emergent baseline VKS scores (Experimental: M = 13.00, SD = 1.69; Control: M = 12.93, SD = 1.28), confirming pre-intervention equivalency. Paired-sample t-tests confirmed significant within-group gains: the Experimental Group experienced a mean gain of 18.60 points (posttest M = 31.60; Cohen's d = 3.18), with most students advancing to Developing and Proficient levels on the VKS, while the Control Group's mean gain of 2.00 points (posttest M = 14.93; d = 0.37) left no Control Group students advancing from the Pre-emergent stage. Posttest between-group comparison yielded $t(28) = 14.87$, $p < .001$, $d = 5.43$, confirming highly significant differences in vocabulary development outcomes between the two groups due to AI-assisted intervention. Therefore, the null hypothesis was rejected. These results provide strong quasi-experimental evidence that teacher-mediated structured AI applications provided significantly greater vocabulary development for Pre-K children in Tacloban City, Philippines than did the traditional instructional approach alone.

Keywords: artificial intelligence, Buddy.ai, early childhood education, Google Read Along, Philippines, pre-kindergarten, quasi-experimental, readability, vocabulary acquisition, Vocabulary Knowledge Scale

The A.I. KID study: Evaluating artificial intelligence interactivity and vocabulary acquisition in pre-kindergarten learners

1. Introduction

The rapid growth of digital technology in early childhood education settings has made artificial intelligence (AI) an increasingly viable instructional strategy for young language learners. In the Philippines, children aged four to five are increasingly exposed to digital devices before entering school, although the majority of these experiences has been passive in nature (e.g., watching animated videos; listening to narrated stories, or tapping a screen to play a game). There is increasing evidence linking passive engagement to limited gains in expressive vocabulary development, reduced ability to engage in depth of conversation, and a lack of growth in reading readiness skills (Chaudron & Di Gioia, 2024). As such, the A.I. KID Study emerged as a direct challenge to this model by investigating whether teacher-guided interaction with AI applications would lead to an increase in vocabulary acquisition among Pre-Kindergarten students in Tacloban City.

The global literacy landscape is dire. According to The World Economic Forum (2024), hundreds of millions of school-aged children around the world do not have the ability to read a simple sentence when they complete primary school. The Programme for International Student Assessment (PISA) scores show a large difference in reading skills for 15 year olds in the Philippines based on PISA's 2022 assessments. This supports evidence that there are some systemic issues (that started with preschool) that are impacting children's literacy skills. Children in the Pre-Kindergarten years (ages approximately 3 - 5) have an excellent opportunity for developing their language skills because, at this age, there is a maximum amount of ability for developing the associated neural pathways used for phonological awareness, semantic memory and syntactic processing (Vygotsky, 1978). Therefore, any efforts made to address literacy acquisition during this time period will yield disproportionate gains in terms of future literacy achievement.

Conversational AI tools developed specifically for young children provide a unique opportunity to assist with closing this achievement gap. Applications such as Buddy.ai, Google Read Along, and Readability provide a gamified, child-friendly environment in which children are not merely observing what happens but actively participating by speaking, responding, and demonstrating comprehension in real time. Furthermore, these platforms are unlike traditional modes of instructional media in that they are able to adapt to the pace of each individual child and provide immediate, corrective feedback while sustaining interest through active participation. Research further suggests that low-anxiety, gamified AI environments promote verbal risk-taking among young learners, leading to increased language production (Hidayat, 2024). UNESCO (2024) is encouraging teachers to assume a digital mentoring role to provide secure, ethical, and pedagogically relevant means of exploring the transformational possibilities of generative AI in early education.

In February 2025, the Department of Education (DepEd) institutionalized the use of AI-assisted education in the Philippines with the opening of the Education Center for AI Research (E-CAIR) by officially incorporating tools such as Buddy.ai and Google Read Along into the basic education system. Based on the findings of Acuña (2025), Filipino children who have utilized gamified AI tools to support their literacy development have shown noticeable improvement in motivation as well as production of language when compared to children instructed through traditional means.

Although the body of evidence supporting the use of AI for learning in primary and secondary schools continues to grow, there is still a significant gap in the empirical literature regarding the use of AI to teach children at the pre-kindergarten level in Tacloban City. There have only been two studies that employed standardized psycholinguistic measures (such as the Vocabulary Knowledge Scale; Wesche & Paribakht, 1996) to measure depth of vocabulary for very young Filipino children. This study aims to fill that gap by attempting to examine if the use

of structured incentives (AI) with a 'teacher as facilitator' results in statistically significantly and practically meaningfully increased vocabulary acquisition in pre-kindergarten children, using an experimental design - two groups (experimental and control) will have a pre-test (before) and post-test (after) for measurement of vocabulary acquisition.

On the Philippine context, two studies have been conducted recently that relate to this topic. Candilas (2025) performed a thematic synthesis of AI applications/uses in language education. Elepaño (2024) studied the effects of gamification upon student performance. Both studies lend some support for increasing technological assistance in the instruction of students in Philippine classrooms; however, neither has focused specifically upon the pre-kindergarten level or have used the Vocabulary Knowledge Scale to evaluate depth of vocabulary. Consequently, there exists a gap in the literature which this study will address.

Objectives of the Study - This study aimed to evaluate the effectiveness of interactive AI technologies at improving the vocabulary skills of Pre-K students in Tacloban City, Philippines. Specifically, this study investigated: (1) the amount of vocabulary that each group had acquired before the intervention; (2) the amount of vocabulary that each group had acquired following the intervention; and (3) whether there was a statistically significant difference in the degree to which each group improved their respective vocabulary skills after the intervention.

Null Hypothesis (H_0) - There is no significant difference in the posttest vocabulary acquisition scores between the Experimental Group and the Control Group of Pre-Kindergarten learners after the utilization of interactive AI applications.

2. Methods

Research Design - A quasi-experimental two-group pretest–posttest design was used in this study. A quasi-experimental design is appropriate when it is not possible to use random assignment due to practical, ethical, or institutional reasons; however, the researcher wishes to provide evidence of comparative causation between the two groups using a non-equivalent control group (Field, 2018). To this end, two intact Pre-Kindergarten class sections—the morning section and the afternoon section—served as the Experimental group (n = 15) and the Control group (n = 15), respectively. During the same two-week period, both groups received conventional Pre-Kindergarten instruction, and the Experimental Group received a ten-session teacher-mediated AI-assisted vocabulary intervention. Both groups were pre-tested and post-tested using the VKS; the near-identical baseline VKS profiles of both groups confirmed equivalence and validity, providing the empirical foundation for valid between-group comparisons.

The design the researcher selected was a concurrent control group because it optimizes internal validity compared to the pre-experimental one group design; thus, any differences in outcomes between the two groups will be less likely attributed to maturation, incidental language exposure or pre-existing familiarity with the English measurement tool than to the effects of the intervention itself. The researcher understands that using intact class sections will limit somewhat the extent to which selection bias can be completely removed; however, the equivalency of pre-test scores and the degree of instructional condition consistency associated with each group significantly reduce the potential effects of this limitation.

Participants - The sample included Pre-K pupils from an early childhood centre located in Tacloban City, Leyte, Philippines during Academic Year 2025–2026. The sample consisted of 30 Pre-K pupils in total (Experimental Group: 15 morning; Control Group: 15 afternoon), ages 3 years and 6 months to 5 years and 11 months. Participants were selected using purposive sampling based on the following inclusion criteria: enrollment in the Pre-Kindergarten program, completion of the pretest, written parental consent, willingness to participate in digital device-related activities, and no current enrollment in Speech-Language Therapy or any other specialized language intervention. Written informed parental consent and child assent were obtained prior to data collection, in compliance with Leyte Normal University ethical standards and Republic Act No. 10173, also known as the

Data Privacy Act (2012).

Instrument - The Vocabulary Knowledge Scale (VKS), a five-level incremental vocabulary assessment tool developed by Wesche and Paribakht (1996), was employed to determine vocabulary depth for two test occasions on 10 target vocabulary words: *curious, enormous, patient, furious, brave, gentle, gloomy, cozy, clever,* and *graceful*. The minimum possible score is 10 and the maximum is 50. The VKS has been recognized as a valid instrument for measuring incremental word knowledge in both first- and second-language contexts (Hadley & Dickinson, 2018). Recent systematic review literature has further provided evidence supporting its validity for indicating vocabulary depth in language learners (Zeng et al., 2025).

Intervention - The program used three different software programs or platforms in the following sequence: 1. Buddy.ai (Version 4.12) for vocabulary activation (approximately eight minutes). 2. Google Read Along (Version 1.64) for guided oral reading (approximately 14 minutes). And 3. Readability for consolidation of comprehension (approximately eight minutes). The 20-20-20 Protocol for eye health was utilized at the midpoint (15 minutes) of the session. The classroom teacher served as the facilitator of the program and provided culturally responsive activities for follow up to student responses generated by the Artificial Intelligence. The control group received instruction from a traditional curriculum and was not provided access to any of the three platforms during the study.

Theoretical Framework - There are four complementary theoretical approaches that serve as the foundation for this theoretical architecture: Mayer's Cognitive Theory of Multimedia Learning (2001), Paivio's Dual Coding Theory (1971), Sweller's Cognitive Load Theory (1988), and Vygotsky's Sociocultural Theory of Cognitive Development (1978). These four theories work together to illustrate how using structured, scaffolded, multimedia artificial intelligence (AI) interaction within the Zone of Proximal Development (ZPD) will produce meaningful vocabulary growth for young children.

Data Analysis - Descriptive statistics of pretest and posttest data were calculated for both groups. For the null hypothesis, three inferential analyses were conducted: paired-sample *t*-tests for each group to compare within-group results, and an independent-samples *t*-test to compare between-group posttest results. Normality was assessed using the Shapiro-Wilk test, and the assumption of homogeneity of variance was checked using Levene's test. Effect size was measured using Cohen's *d* (Cohen, 1988). All analyses were completed using IBM SPSS Statistics, version 28, at $\alpha = .05$, two-tailed.

3. Results and Discussion

3.1 Pretest Vocabulary Profiles: Experimental and Control Groups

Table 1
Pretest VKS Scores of the Experimental Group and Control Group (N = 15 per group)

VKS Score Descriptor	Range	Exp. f	Exp. %	Ctrl. f	Ctrl. %	Interpretation
Pre-emergent	10–18	15	100.0%	15	100.0%	No/minimal word knowledge
Emerging	19–26	0	0.0%	0	0.0%	Recognizes words; limited meaning
Developing	27–34	0	0.0%	0	0.0%	Functional understanding
Proficient	35–42	0	0.0%	0	0.0%	Uses words in context
Advanced	43–50	0	0.0%	0	0.0%	Full mastery
Mean (SD)	–	13.00 (*SD* = 1.69)	12.93 (*SD* = 1.28)			Pre-emergent (both)

Note. VKS = Vocabulary Knowledge Scale (Wesche & Paribakht, 1996). *Exp.* = Experimental Group (morning class, $n = 15$); *Ctrl.* = Control Group (afternoon class, $n = 15$). Scores range from 10 to 50. *f* = frequency; % = percentage within group. *SD* = standard deviation.

Table 1 shows the VKS pretest scores of both groups prior to the intervention. The Experimental Group had a

pretest mean of 13.00 (*SD* = 1.69) and the Control Group had a pretest mean of 12.93 (*SD* = 1.28). All 30 participants across both groups scored within the Pre-emergent range (10–18), indicating no prior knowledge of the 10 target vocabulary words. The near-identical pretest means confirm baseline equivalency between the groups, which is a critical prerequisite for valid quasi-experimental comparison. This Pre-emergent profile is consistent with the developmental patterns observed among Pre-Kindergarten children in the Philippines, where early vocabulary exposure has been predominantly passive in nature (Chaudron & Di Gioia, 2024).

3.2 Posttest Vocabulary Profiles: Experimental and Control Groups

Table 2
Posttest VKS Scores of the Experimental Group and Control Group (N = 15 per group)

VKS Score Descriptor	Range	Exp. f	Exp. %	Ctrl. f	Ctrl. %	Interpretation
Pre-emergent	10–18	0	0.0%	15	100.0%	No/minimal word knowledge
Emerging	19–26	0	0.0%	0	0.0%	Recognizes words; limited meaning
Developing	27–34	12	80.0%	0	0.0%	Functional understanding
Proficient	35–42	3	20.0%	0	0.0%	Uses words in context
Advanced	43–50	0	0.0%	0	0.0%	Full mastery
Mean (SD)	–	31.60 (*SD* = 2.87)	14.93 (*SD* = 1.28)			Developing vs. Pre-emergent

Note. VKS = Vocabulary Knowledge Scale (Wesche & Paribakht, 1996). *Exp.* = Experimental Group (morning class, *n* = 15); *Ctrl.* = Control Group (afternoon class, *n* = 15). Posttest administered within one week of the final session using randomized word order.

Table 2 shows a significant difference between both groups in the post-test results. The mean Post-Test score for the Experimental group was 31.60 (SD = 2.87) representing an 18.60 point gain from Baseline. Out of the total 15 children in the Experimental group at the time of the post-test, 12 (80%) had progressed to the Developing level, and 3 (20%) had progressed to the Proficient level. The Control group had a mean Post-Test score of 14.93 (SD = 1.28) with only a 2.00 point gain from Baseline and no children progressed to beyond the Pre-emergent level at the time of the post-test. Qualitative analysis of the session logs showed that children in the Experimental Group spontaneously began producing some of the target vocabulary words (e.g., gentle, gloomy, and cozy) without prompting within 2 weeks of the introduction of those words, while this was not the case in the Control Group. The data provide additional evidence that vocabulary development was determined by the use of AI-mediated interactivity, rather than through incidental exposure to vocabulary in a classroom setting.

3.3 Inferential Analysis: Within-Group and Between-Group Comparisons

Table 3
Within-Group Paired-Sample t-Test Results and Between-Group Independent-Samples t-Test (N = 15 per group)

Comparison	Mean Diff.	*t*	*df*	*p*	Cohen's *d*
Experimental Group: Pre vs. Post (Paired t-test)	18.60	22.14	14	< .001	3.18
Control Group: Pre vs. Post (Paired t-test)	2.00	2.31	14	.037	0.37
Between Groups: Experimental vs. Control Posttest (Independent t-test)	16.67	14.87	28	< .001	5.43

Note. Paired-sample *t*-tests used for within-group comparisons. Independent-samples *t*-test used for between-group posttest comparison. The independent samples T test for Experimental Group (M = 31.60) and Control Group (M = 14.93) indicates an unusually large effect size (t(28) = 14.87, p < .001, d = 5.43), thereby providing convincing evidence that a highly significant difference existed in vocabulary acquisition between the two groups.

This means that vocabulary acquisition of the Experimental Group utilizing AI-assisted instruction had an unusually high level of success compared to the Control Group; therefore, the null hypothesis was rejected. These results support the theoretical framework of the study. The vocabulary gains of the Experimental Group are consistent with Mayer's (2001) dual-channel multimedia engagement, Paivio's (1971) dual encoding of verbal and

imagistic memory traces, Sweller's (1988) management of cognitive load through scaffolding, and Vygotsky's (1978) mediated scaffolding within each child's Zone of Proximal Development. Evidence is also provided by Nguyen and Zisook (2025), who showed that vocabulary acquisition increases rapidly for emerging readers who receive timely, accurate feedback from an AI tutoring system. The immediate feedback loops afforded by Google Read Along's real-time corrective feedback and Buddy.ai's adaptive conversational scaffolding induced accelerated vocabulary acquisition for the Experimental Group in this study.

4. Conclusion

The research comprised a quasi-experimental two-group pretest–posttest methodology that aimed to ascertain the degree to which structured, teacher-facilitated AI-supported instruction provides an avenue for improved vocabulary acquisition for Pre-Kindergarten children in Tacloban City. Both groups demonstrated similar Pre-emergent VKS profiles (Experimental: $M = 13.00$; Control: $M = 12.93$) prior to the study. During the two weeks of intervention, the Experimental Group had a significantly greater increase in vocabulary ($M = 31.60$, $d = 3.18$) than the Control Group ($M = 14.93$, $d = .37$). A comparison of the two groups confirmed that the Experimental Group had a significantly greater increase in vocabulary than the Control Group ($t(28) = 14.87$, $p < 0.001$, $d = 5.43$), thus rejecting the null hypothesis. The results of the A.I. KID Study provide strong evidence from a quasi-experimental study demonstrating that Pre-Kindergarten students who received structured and teacher-facilitated interactions with the use of Buddy.ai, Google Read Along and Readability, made significantly greater gains in vocabulary after two weeks than those who only received traditional instruction. These findings also support the use of AI tools as evidence-based resources in the implementation of basic education in the Philippines, consistent with the Department of Education's E-CAIR Initiative.

Practical Educational Implications - The results of the AI KIDS Study provide valuable insights for practitioners (teachers), students, and schools within the context of the Philippines' educational system.

For **teachers**, according to the research, teachers can achieve large vocabulary gains by injecting a structured three-platform A.I. learning session into their daily Pre-Kindergarten literacy program —along with a 20-20-20 eye health protocol — in as little as two weeks, even without having expertise with A.I., when acting as a culturally responsive facilitator in conjunction with A.I. tools. It is recommended that teachers should receive professional development in digital mentoring to support effective A.I. implementation, as is supported by UNESCO (2024).

For **students**, the finding of this study indicates that, for students, interactive A.I. learning environments access to responsive, gamified A.I. applications may accelerate vocabulary acquisition and reading readiness for Pre-Kindergarten students who, prior to school, have primarily been exposed to passive digital content. Furthermore, they can aid in the development of vocabulary depth and reading readiness during the critical window of opportunity identified by Vygotsky (1978).

For **schools and administrators**, the findings of the study provide schools and administrators with evidence of the effectiveness of implementing AI-assisted vocabulary instruction for all Pre-Kindergarten students in the context of the DepEd's E-CAIR initiative. Schools are encouraged to routinely invest in age-appropriate A.I. platforms for early childhood education and to establish policies for responsible and supervised A.I. usage in early childhood settings.

Limitations - One of the major limitations of this research is that the participants were not randomly assigned to groups—intact classes were used. The research was carried out in a single private school located in Tacloban City, limiting generalizability of the findings. The intervention period was only two weeks, and there was no delayed retention test; and three separate AI applications were evaluated concurrently, making it impossible to determine the independent contribution of each application to the outcome measures.

Recommendations - Future researchers are encouraged to repeat this study using true experimental, delay post-test, and that include multiple public and multi-language settings, in addition to using independent blind raters,

and also apply component analysis designs and mixed methods approaches, in order to isolate the effects of each AI tool.

AI Use Disclosure - Claude (Anthropic, version claude-sonnet-4-20250514; 2025) was used as a language editing and formatting tool for my/our manuscript. All outputs were verified, reviewed, and edited by the author/s to ensure fact-checking, statistical validity, and citation accuracy. No personal or identifying information was submitted to AI. The author/s assume sole responsibility for this work..

5. References

Acuña, R. M. (2025). Artificial intelligence in Philippine education: A narrative review of applications, perceptions, and challenges. *International Research Journal of Science, Technology, Education, and Management, 5*(2), 25–38.

Buddy.ai. (2026). *Buddy.ai: The mobile AI tutor for kids* (Version 4.12) [Mobile app]. Google Play Store.

Candilas, K. (2025). The permeates of artificial intelligence (AI) in language education: A thematic synthesis. *Journal of Interdisciplinary Perspectives, 3*(3), 115–123.

Chaudron, S., & Di Gioia, R. (2024). *The future of child development in the AI era: Implications for cognition and socio-emotional skills*. Everyone.AI Research Paper.

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

Department of Education. (2025, February 20). *DepEd launches Education Center for AI Research (E-CAIR) to drive innovation in Philippine basic education*.

Elepaño, C. J. T. (2024). *The effect of gamification on the self-directed learning skills and academic performance of students* [Master's thesis, De La Salle University]. Animo Repository.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Google. (2026). *Read Along by Google: A fun reading app for kids* (Version 1.64) [Mobile app]. Google Play Store.

Hadley, E. B., & Dickinson, D. K. (2018). Measuring young children's word knowledge: A conceptual review. *Journal of Early Childhood Literacy, 20*(4), 714–743.

Hidayat, A. (2024). Low-anxiety environments: The impact of gamified AI on verbal risk-taking in early childhood education. *Global Journal of Educational Technology, 12*(2), 88–101.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.

Nguyen, T., & Zisook, S. (2025). Scaffolded vocabulary growth: Immediate feedback loops in AI-driven tutoring systems. *Thai Nguyen University Journal of Science and Technology*.

Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart, and Winston.

Republic Act No. 10173. (2012). *An act protecting individual personal information in information and communications systems in the government and the private sector*. Congress of the Philippines.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285.

UNESCO. (2024). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth vs. breadth. *Canadian Modern Language Review, 53*(1), 13–40.

World Economic Forum. (2024). *Shaping the future of learning: The role of AI in education 4.0*. World Economic Forum.

Zeng, Y., Kuo, L.-J., Chen, L., Lin, J.-A., & Shen, H. (2025). Vocabulary instruction for English learners: A systematic review connecting theories, research, and practices. *Education Sciences, 15*(3), 262.