InvestiGO
Revista Científica Multidisciplinaria
ISSN 2953-6367
http://revistainvestigo.com

Octubre 2026
Vol. 7 No. 21, PP. 50-63
https://doi.org/10.56519/4c0d0b95

APRENDIZAJE MÓVIL DE VOCABULARIO EN INGLÉS CON LINGODEER EN UNA ESCUELA RURAL DEL VALLE DEL CHOTA, ECUADOR: ESTUDIO PRIMARIO PRETEST-POSTEST

MOBILE VOCABULARY LEARNING WITH LINGODEER IN A RURAL SCHOOL OF THE CHOTA VALLEY, ECUADOR: A PRIMARY PRETEST-POSTTEST

Carlos Leomar De La Cruz Espinoza[1], Martha Viveros Almeida[2]

{carlos.delacruz@upec.edu.ec[1], martha.viveros@upec.edu.ec[2]}

Fecha de recepción: 29/07/2026 / Fecha de aceptación: 21/08/2026 / Fecha de publicación: 14/10/2026

RESUMEN: En las escuelas rurales, las limitadas oportunidades de exposición al inglés y de práctica autónoma del vocabulario justifican la exploración de herramientas móviles integradas a la enseñanza. En este contexto, este estudio cuantitativo primario examinó el cambio en el logro de vocabulario en inglés durante un periodo de enseñanza apoyada por LingoDeer en una escuela pública rural del Valle del Chota, Ecuador. Este estudio cuantitativo primario examinó el cambio en el logro de vocabulario en inglés durante un periodo de enseñanza apoyada por LingoDeer en una escuela pública rural del Valle del Chota, Ecuador. Se analizaron las respuestas a 30 reactivos y los puntajes emparejados de 70 estudiantes, así como un cuestionario diagnóstico de 71 respondientes. Los totales se verificaron reactivo a reactivo. La media aumentó de 12,26 (DE = 5,12) a 25,80 (DE = 1,42) puntos sobre 30; la ganancia media fue 13,54 (DE = 5,22), t(69) = 21,69, p < 0,001, IC 95 % [12,30; 14,79], r pre-post = 0,067 y dz = 2,59. Los 70 estudiantes mostraron diferencias positivas (rango 2-23; prueba exacta de signos, p < 0,001), con una proporción de 7,1 % en AAR o DAR, que se incrementó a 100,0 %. En el cuestionario, las respuestas «Nunca» oscilaron entre 45,1 % y 67,6 % para mediación digital y entre 56,3 % y 63,4 % para desempeño léxico autopercibido. La consistencia interna fue alta para mediación digital (alfa = 0,966) y desempeño léxico (alfa = 0,968), mientras que el postest de logro produjo un coeficiente negativo (alfa = -0,395). La misma prueba se administró en ambas ocasiones, los puntajes finales estuvieron próximos al techo y el archivo analítico no conservó los identificadores de condición ni medidas verificadas de exposición. Por ello, el artículo informa un cambio primario pretest-postest, pero no atribuye causalmente ese cambio a LingoDeer.

Palabras clave: aprendizaje móvil de idiomas; LingoDeer; vocabulario en inglés; TPACK; educación rural; Ecuador

[1]Unidad Educativa Guardiana De Los Saberes San Gabriel de Piquiucho: Docente de Inglés, https://orcid.org/0009-0001-1171-8451; +593981172433.
[2]Universidad Politécnica Estatal del Carchi: Profesor Ocasional, https://orcid.org/0000-0002-2441-6954, +593979289520.

ABSTRACT: In rural schools, limited opportunities for exposure to English and independent vocabulary practice justify exploring mobile tools integrated into teaching. In this context, this primary quantitative study examined the change in English vocabulary achievement during a period of instruction supported by LingoDeer in a rural public school in the Chota Valley, Ecuador. This primary quantitative study examined change in English-vocabulary achievement during a period of LingoDeer supported instruction in a rural public school in Ecuador's Chota Valley. Thirty item responses and matched scores from 70 students were analysed, together with a diagnostic questionnaire completed by 71 respondents. Item totals were verified item by item. Mean achievement increased from 12.26 (SD = 5.12) to 25.80 (SD = 1.42) out of 30; mean change was 13.54 (SD = 5.22), t(69) = 21.69, p < .001, 95% CI [12.30, 14.79], pre-post r = .067, and dz = 2.59. All 70 students had positive differences (range 2-23; exact sign test, p < .001), and the proportion in AAR or DAR increased from 7.1% to 100.0%. For the questionnaire, 'Never' responses ranged from 45.1% to 67.6% for digital mediation and from 56.3% to 63.4% for self-perceived lexical performance. Internal consistency was high for digital mediation (alpha = .966) and lexical performance (alpha = .968), whereas the posttest achievement items produced a negative coefficient (alpha = -.395). The same test was administered at both occasions, final scores were near the ceiling, and the analytic file did not retain condition identifiers or verified exposure measures. The article therefore reports primary pretest-posttest change without attributing that change causally to LingoDeer.

Keywords: mobile-assisted language learning; LingoDeer; English vocabulary; TPACK; rural education; Ecuador

INTRODUCCIÓN

Vocabulary knowledge is the foundation for comprehension and production in listening, speaking, reading and writing. Lexical knowledge is not available for communication until learners have repeated experiences with the form, meaning, pronunciation, and contextual use (1-4). In contemporary research, word knowledge is considered as a multidimensional phenomenon and is differentiated from the productive mastery needed for communication (1,3). The repeated presentation of word lists with spaced intervals of time and attention and retrieval is necessary for acquisition; the presentation of word lists alone is not sufficient for the acquisition of memory (2,4). In a low-proficiency English as a foreign language (EFL) classroom, the limited amount of time for class instruction and the few opportunities for individualized practice can result in a classroom that produces receptive fragments that are not easily transferred to the classroom's oral and written environment.

The present study is based on two theoretical strands. Well-sequenced applications operationalize this principle, which is known as the input-oriented approach, by systematically sequencing and recycling input that is just a little above the learner's level. The sociocultural view of learning sees the zone of proximal development as the space in which learning occurs as a result of mediation (6), which can be provided by software, but requires teacher orchestration to transform individual practice into shared communicative competence.

Mobile-assisted language learning (MALL) can provide portable, repeatable, multimedia practice in the classroom. The recent reviews and meta-analyses, in general, suggest positive results for second-language learning and, in particular, for vocabulary (7-13), although the results are also characterized by significant differences in the effects across learners' ages, intervention duration, learning design, and assessment. Mobile practice as a pedagogical tool is not a treatment itself, but rather a pedagogical tool, and these syntheses do not exclude the possibility of publication bias, weak comparison conditions or incomplete reporting, which can inflate effect estimates. Convenience and gamified feedback are added to the commercial applications, but the mere presence of an application does not guarantee learning. Some of the concerns that persist are small samples, brief exposure, attrition, novelty effects, and weak comparisons (14,15). Incomplete reporting of statistical assumptions, effect sizes, measurement reliability and validity also are common issues identified in reviews of quantitative MALL research (16,17).

The field has evolved from feasibility reports to controlled evaluations. An early synthesis revealed mostly positive results across implementations (18), and more recent meta-analyses indicate moderate to large average effects across implementations, with significant heterogeneity, but not effects similar to the corrected within-person dz reported here (9,10). The proposed mechanisms are: more time on task, immediate corrective feedback, multimedia coding, and spaced review (19). The evidence at the platform level is still inconsistent and is contingent on persistence, dosage, learning design and comparison quality (15). Therefore, a large pooled pre-post standardized change should not be compared directly or interpreted as application efficacy.

LingoDeer organizes each language course into a set of language units that are fixed in terms of difficulty level and content from beginner to intermediate. Each unit includes explicit grammar notes, presentation of words in context, native-speaker recordings, immediate item-level feedback and progress analytics, as well as a dedicated spaced-repetition review module (21). This is a structured, explanation-rich profile and is not as much an implicit, streak-driven platform as it is a platform that might be articulated in the classroom, as unit topics can be mapped onto national curriculum blocks. However, empirical assessment is still limited. A recent qualitative case study reported that Indonesian junior secondary learners showed greater interest, involvement and reported word recall, while the difficulties of pronunciation and cost of premium version were reported (22). The present study fills in this gap in quantitative achievement evidence and in evidence from Latin America.

These concerns are more significant in rural public-school environments where access to devices, connectivity, and technical support can be inconsistent. The Ecuadorian EFL curriculum is structured around five strands of learning (communication and cultural awareness, oral communication, reading, writing, and language through the arts) and has progressive expectations across General Basic Education (23) that align with the Common European Framework of Reference for Languages (24). The most defensible use of digital practice is when it is used to fulfill these curricular goals and is mediated in a purposeful way by the teacher.

The technological pedagogical content knowledge (TPACK) framework supplies that rationale. TPACK treats effective technology integration as the situated interaction of content,

pedagogical, and technological knowledge rather than as the isolated use of a digital tool (25,26). Applied to MALL, TPACK implies that an application's affordances produce learning only when the teacher aligns them with content standards and appropriate pedagogy. In the present context, LingoDeer supplied structured multimedia exercises and immediate feedback, while the teacher connected those activities to curricular vocabulary, the ERCA cycle (experience, reflection, conceptualization, and application), and communicative classroom tasks.

This adaptation was especially important in the predominantly Afro-Ecuadorian rural community of the Chota Valley where opportunities to use English outside of school may be limited, and devices, connectivity, and technical support may be uneven. As a result, LingoDeer was designed to be a teacher-led, curriculum-based tool, not a standalone application that requires constant internet connectivity.

The research team conducted the study in a public educational unit in Piquiucho, a predominantly Afro-Ecuadorian community in the Chota Valley of Carchi Province. Classroom control and intervention assignments were planned, but the final participant-level analytic file retained grade only (8, 9, or 10), not class section or condition. The article thus reports a primary pooled pre test-post test study as opposed to a controlled treatment-effect estimate. Three research questions were addressed: (a) What were baseline patterns of teacher digital mediation and self-reported lexical performance? (b) What were pooled and grade-specific vocabulary-score changes from pretest to posttest? and (c) What were the measurement and design constraints for the interpretation of those changes?

MATERIALES Y MÉTODOS

Design, setting, and data collection. The present article presents the main quantitative one-group pretest-posttest analysis carried out at the Guardiana de los Saberes San Gabriel de Piquiucho Educational Unit. This study used a baseline contextual questionnaire and a 30-item vocabulary assessment, which were given before and after a period of instruction supported by LingoDeer. The research team documented the responses for each item on the instrument at the participant level and matched the achievement records with coded identifiers. These basic study records were used to calculate all numerical results presented in this article. No qualitative data were gathered or analysed.

Participants and analytic samples. There were 70 students who completed the matched 30-item pretest and posttest responses for the achievement analysis, and 71 respondents who completed the contextual questionnaire. The one-participant difference was due to the absence of a student who completed the contextual questionnaire on the day of the achievement posttest, so that student could not be included in the matched pretest-posttest analysis.

The 70 complete achievement pairs were included, including four students who were identified as having special educational needs (SEN) in the enrolment summary. They were assessed as part of the school's regular inclusive classroom practices and the assessment content and scoring criteria were not changed so that the assessment was comparable. There

was no SEN based exclusion or score adjustment. Unfortunately, individual SEN status and specific support measures were not stored in the participant-level analytic file, meaning that SEN-specific sub-group analyses were not possible; this is acknowledged.

Table 1. Participant distribution, curricular target, and planned classroom assignment.

Grade / class group	Boys	Girls	Total	SEN	Planned classroom role / analytic status
Eighth A	5	5	10	2	Planned control; section/condition not retained in analytic file
Eighth B	5	5	10	2	Planned intervention; section/condition not retained in analytic file
Ninth A	4	6	10	0	Planned control; section/condition not retained in analytic file
Ninth B	5	5	10	0	Planned intervention; section/condition not retained in analytic file
Tenth (single group)	12	18	30	0	LingoDeer-supported instruction; no comparator
Total	**31**	**39**	**70**	**4**	**70 complete pairs included; no participant-level condition or SEN field**

Note. Planned classroom roles, class sections and SEN are recorded as part of the enrollment records. The participant-level analytic file includes coded ID, grade, item responses, and total score but lacks section, condition, sex, and SEN fields. All 70 complete pairs were included, so there is no possibility of separating the planned control and intervention classrooms for a between groups analysis that is reproducible.

Instruments. A locally developed objective test with 30 items and a scoring of 1 point for each correct answer was used as the achievement instrument (0-30). The same form was administered at pretest and posttest; no parallel form or delayed retention test was used. Expert content review and alignment with CEFR A1-A2 and the Ecuadorian curriculum were reported, but independent criterion validation was not conducted. Analysis of the primary binary item responses gave alpha/KR-20 = .814 at pretest and alpha = -.395 at posttest. The negative posttest coefficient is a negative average inter-item covariance and is therefore not a valid estimate of internal consistency for this administration. Unstable inter-item covariances may have been due to the ceiling effect (M = 25.80/30, SD = 1.42) and the lack of variance in scores. The total scores for the post test were therefore considered exploratory. The diagnostic questionnaire comprised 14 items: 1-8 were related to teacher digital mediation, 9-13 were related to self-perceived lexical performance and Item 14 was related to resource preference. Alpha values were .966 for Items 1-8 and .968 for Items 9-13, with a 13-item alpha of 0.983. These coefficients describe internal consistency only and do not establish unidimensionality or construct validity.

Table 2. Measurement framework and data-collection instruments.

Instrument	Construct and content	Scale	Quality evidence reported
Achievement pre-/post-test	Meaning recognition; contextual vocabulary; simple-sentence production; auditory discrimination	30 items; 0-30	Primary item-level responses; pretest alpha/KR-20 = .814; posttest alpha = -.395 (not acceptable)
Diagnostic questionnaire	Teacher digital mediation (Items 1-8); self-perceived lexical performance (Items 9-13)	4-point frequency scale (13 items)	Items 1-8 alpha = .966; Items 9-13 alpha = .968; all 13 items alpha = .983; factor structure untested
Resource-preference item	Preferred resource for vocabulary development (Item 14)	Single categorical choice	Audio was the modal response: 35/71 (49.3%)

Note. Resource-preference results were calculated from the 71 primary questionnaire responses: audio 35 (49.3%), reading 19 (26.8%), dialogues 13 (18.3%), and dictation 4 (5.6%).

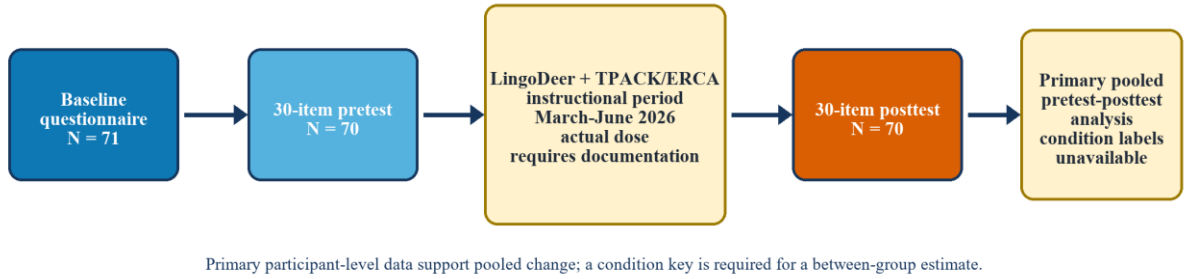


Figure 1. Study procedure and analytical scope.

Note. The diagram presents the primary-data study sequence and the pooled analytic scope. A condition key and documented exposure are required before a control-versus-intervention effect can be estimated.

Table 3. Curriculum-to-activity alignment in the instructional guide.

Curriculum thread	Grade 8 (A1.1)	Grade 9 (A1.2)	Grade 10 (A2.1)
Communication and cultural awareness	Greetings and introductions	Food origins and healthy choices	Months, festivals, and birthdays
Oral communication	Family dialogues	Animals and sound-based practice	Buying clothes: buyer-seller exchanges
Reading	Short descriptions of pets	Menus and food labels	Schedules, time, and agendas
Writing	Classroom objects and demonstratives	Routines and short messages	Personal diary entries
Language through the arts	Colours and visual labelling	Weather vocabulary	Seasons and comparative descriptions

Note. Levels and content follow the study's curriculum-convergence matrix and instructional units. A1.1, A1.2, and A2.1 denote the CEFR sublevels set by the Ecuadorian curriculum for eighth, ninth, and tenth grade respectively.

Statistical analysis. Primary item totals were calculated for every participant and matched by coded ID across occasions. Integer totals were preserved and summary values were rounded only after calculation. Means, sample standard deviations, standard errors, ranges, and achievement-band frequencies were computed from the 70 complete paired records. For the paired-samples t test, the pretest-posttest data and a two-sided alpha of .05 were used, and for the 95% confidence interval, df = 69. Participant-level scores were used to calculate Pearson's pre-post correlation, Cohen's dz, and Glass's delta. The paired-samples t test was used and supplemented by an exact two-sided sign test as a sensitivity analysis, due to the questionable normality of the paired difference scores and substantial grade heterogeneity. The item-level responses were used to compute alpha/KR-20 coefficients for the achievement test and for the two subscores of the questionnaire. The analytic file does not contain any condition identifiers, so an effect of control versus intervention was not estimated. Achievement bands were DAR 27-30, AAR 21-26, PAAR 13-20, and NAR 0-12.

RESULTADOS

Baseline digital mediation. Primary questionnaire responses (N = 71) showed that digital tools were seldom integrated into English instruction. 'Never' responses were 57.7% for digital applications as class materials, 62.0% for English audio through applications, 45.1% for audio-repetition exercises, 50.7% for reading-comprehension application frequency, 62.0% for their actual use, 57.7% for writing exercises supported by applications, 63.4% for online applications supporting English learning, and 67.6% for pronunciation practice. Every percentage is an attainable fraction of 71 and was calculated from its count.

Self-reported lexical performance. 'Never' responses were 57.7% for understanding main ideas when reading, 63.4% for understanding vocabulary in audio input, 60.6% for using vocabulary in everyday speech, 56.3% for understanding vocabulary while reading, and 63.4% for using appropriate vocabulary in writing. Audio was the modal preferred resource (35/71, 49.3%), followed by reading (19/71, 26.8%), dialogues (13/71, 18.3%), and dictation (4/71, 5.6%). These findings answer the first research question as baseline context; they are not presented as outcomes of the intervention.

Questionnaire reliability and response pattern. The alpha for the teacher digital mediation (Items 1-8) was .966, for self-perceived lexical performance (Items 9-13) was .968, and for all 13 scaled items, alpha was .983. Separate subscale coefficients prevent the two conceptually different domains from being considered as a single construct. Very high alpha values, however, may be due to redundancy in the items or to a highly patterned response matrix and are not indicative of factorial validity. Factor analysis, item-total diagnostics, test-retest reliability, and external validation were not conducted.

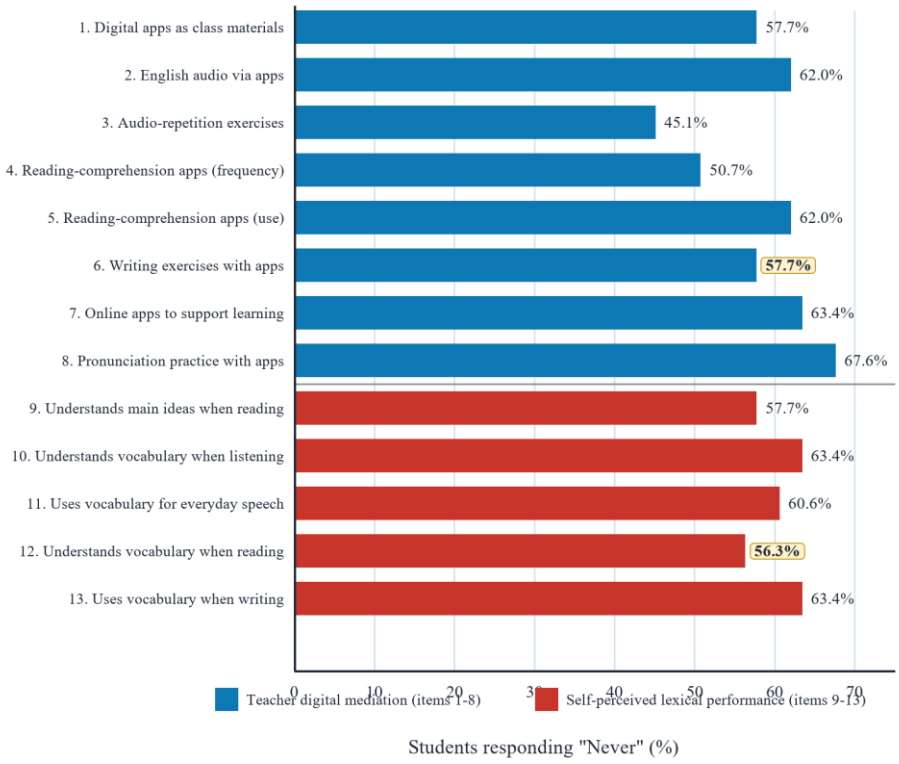


Figure 2. Baseline «Never» responses for digital mediation and lexical performance.

Note. N = 71. Percentages were calculated from primary questionnaire counts. The two corrected rounding values, 41/71 = 57.7% and 40/71 = 56.3%, are highlighted in yellow in the figure.

Pooled vocabulary-test change. For the 70 matched records, the pretest total was 858 (M = 12.26, SD = 5.12, range 5-23) and the posttest total was 1806 (M = 25.80, SD = 1.42, range 23-29). The integer paired-difference total was 948, giving a mean change of 13.54 (SD = 5.22, SE = 0.62, range 2-23), t(69) = 21.69, p < .001, 95% CI [12.30, 14.79]. The pre-post correlation was r = .067, Cohen's dz = 2.59, and Glass's delta = 2.64. All 70 paired differences were positive (range 2–23); the two-sided sign test was exact, p < .001, and did not assume normally distributed paired differences. The sign test is used to assess the direction of change, not the size of the mean difference, and thus complements (not replaces) the paired t test.

Table 4. Pooled pretest-posttest vocabulary results.

Outcome	N	Mean	SD	SE	Difference / inferential result
Pretest	70	12.26	5.12	0.61	n/a
Posttest	70	25.80	1.42	0.17	n/a
Posttest - pretest	70	13.54	5.22	0.62	t(69) = 21.69; p < .001; 95% CI [12.30, 14.79]
Standardized / sensitivity	70	n/a	n/a	n/a	dz = 2.59; Glass delta = 2.64; r = .067; exact sign p < .001

Note. Scores range from 0 to 30. All values were calculated from 70 primary participant-level records; totals were integers and means were rounded to two decimals only after calculation. The exact mean difference is 948/70 = 13.542857. dz standardizes the paired difference; Glass's delta uses the pretest SD. Neither is a between-group causal effect.

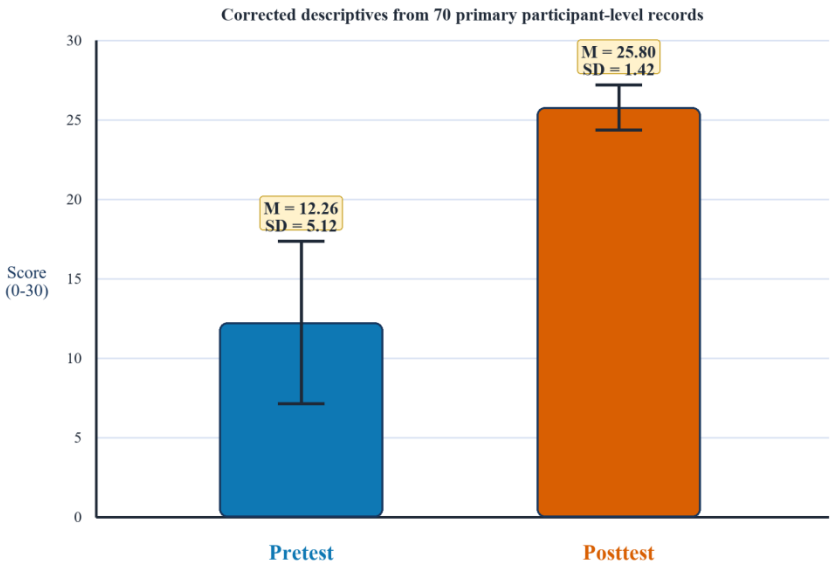


Figure 3. Mean vocabulary scores before and after the instructional period.

Note. Error bars represent +/- 1 sample SD. Yellow labels identify the corrected values calculated from the primary participant-level data.

Achievement-band distribution. At pretest, 42 students (60.0%) were in NAR, 23 (32.9%) in PAAR, 5 (7.1%) in AAR, and none in DAR. At posttest, 48 (68.6%) were in AAR and 22 (31.4%) in DAR; no participant was in PAAR or NAR. Thus, the combined AAR/DAR proportion increased from 7.1% to 100.0%. These counts were calculated directly from the primary participant-level totals. Without retained condition identifiers, the distribution cannot establish a LingoDeer treatment effect.

Table 5. Distribution across nationally adapted achievement bands.

Achievement band	Code	Score range	Pretest n (%)	Posttest n (%)
Masters the required learning	DAR	27-30	0 (0.0)	22 (31.4)
Achieves the required learning	AAR	21-26	5 (7.1)	48 (68.6)
Close to achieving the required learning	PAAR	13-20	23 (32.9)	0 (0.0)
Does not achieve the required learning	NAR	0-12	42 (60.0)	0 (0.0)
Total	**n/a**	**0-30**	**70 (100.0)**	**70 (100.0)**

Note. The four category labels are translated from the Ecuadorian qualitative scale and proportionally adapted to a 30-point test. DAR = domina los aprendizajes requeridos; AAR = alcanza los aprendizajes requeridos; PAAR = está próximo a alcanzar los aprendizajes requeridos; NAR = no alcanza los aprendizajes requeridos.

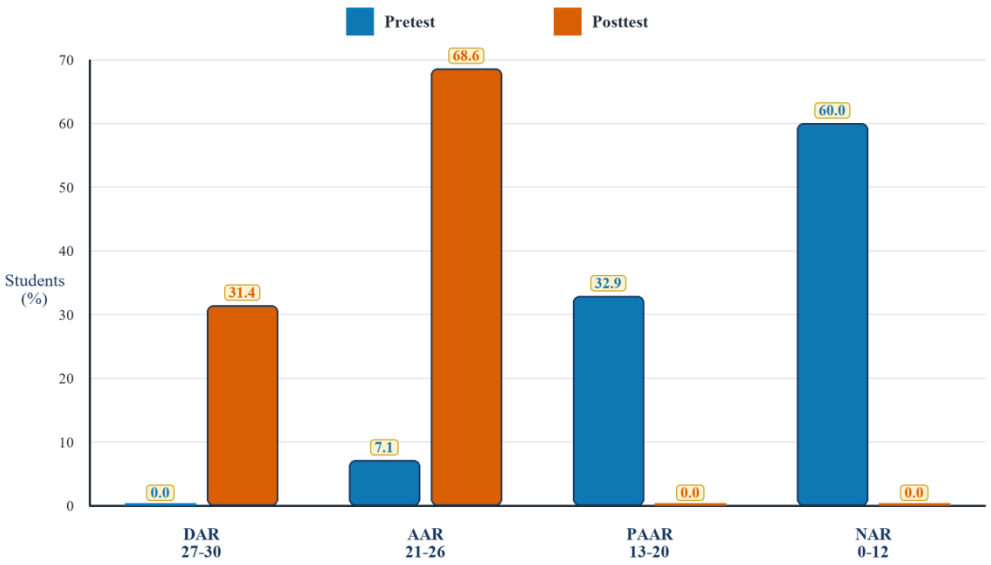


Figure 4. Pretest and posttest achievement-band distributions.

Note. N = 70. Values are percentages calculated from the primary participant-level totals. Yellow labels identify corrected values. DAR, AAR, PAAR, and NAR are defined in Table 5.

Magnitude, heterogeneity, and ceiling. The within-person standardized mean change (dz) is the mean paired difference (13.54) divided by the standard deviation of the paired difference (5.22). It is not an independent-groups effect and should not be interpreted as or compared directly with Hedges' g. The final score compression (restricted posttest variance, SD = 1.42) and the near-ceiling posttest mean (M = 25.80/30), coupled with the large value, suggest that caution should be used when interpreting the results. Therefore, dz is not reported as a large effect size between groups due to LingoDeer. Mean gains differed sharply by grade: 19.20 points in Grade 8 (SD = 2.55), 15.35 in Grade 9 (SD = 1.57), and 8.57 in Grade 10 (SD = 3.02). These differences track distinct baseline levels and show that the pooled statistic mixes grade and curriculum effects. The near-zero pre-post correlation (r = .067) shows little rank-order stability. At posttest every score lay between 23 and 29, the mean was 86% of the maximum, and alpha was negative. Ceiling compression, repeated exposure to the same form, close alignment between instruction and test content, and possible teaching to the test are therefore plausible alternatives to a treatment interpretation.

Summary of findings. The primary records support three reproducible descriptions: limited baseline digital mediation and low self-reported lexical performance, a large pooled increase in integer achievement totals, and strong grade heterogeneity followed by a compressed posttest distribution. The final analytic file does not support a control-versus-intervention comparison because condition membership is absent. The negative posttest reliability coefficient, use of the same form twice, lack of a delayed test, missing actual-dose and fidelity logs, and lack of participant-level SEN identifiers materially limit interpretation. The corrected statistics establish primary pretest-posttest change but do not by themselves establish causality.

DISCUSIÓN

Three main findings emerged from this primary study. First, pre-questionnaire data show that digital mediation is scarce and self-perceived lexical performance is low. Second, the participant-level achievement data reveal a large pooled score gain (all 70 differences were positive). Third, the increase is quite different across grade levels and tapers off to a low, near-ceiling posttest distribution that has negative item reliability. The correct interpretation is a main descriptive change in the teaching period, not that LingoDeer brought about the change.

The baseline results clarify why mobile support was pedagogically relevant. Across the eight digital-mediation items, 45.1%-67.6% selected 'Never'; across the five lexical self-report items, 56.3%-63.4% selected 'Never'. Audio was the modal resource preference. These attainable percentages and separate subscale reliabilities answer the contextual research question without being used as post-intervention evidence.

The corrected mean gain remains much larger than typical educational intervention effects, but several features can inflate it: low and grade-stratified baselines, direct alignment between practised and tested content, immediate feedback, administration of the identical instrument twice, and a posttest ceiling. The fall in SD from 5.12 to 1.42 is descriptive evidence of restricted final-score variance, not proof of equalized learning. The negative posttest alpha further cautions against interpreting the total as a stable latent measure.

TPACK focuses on how LingoDeer contributes to the practice rather than the practice of LingoDeer itself, that is, teacher-led pedagogical orchestration. The teacher chose the technological affordances of the application to achieve specific lexical and phonological goals, integrated them in the ERCA cycle (experience, reflection, conceptualization, application), and linked the digital practice with communicative follow-up exercises that were related to the grade level curriculum. Therefore, the likely teaching process was the teacher's alignment and mediation: LingoDeer provided structured multimedia practice and immediate feedback, and the teacher was able to use these to create contextualized learning opportunities that were linked to the five curricular strands and the progression from A1.1 to A2.1 (23,24). Note that this interpretation should not be understood as evidence that LingoDeer or teacher mediation alone caused the observed change because the study design did not isolate these components experimentally.

The outcome is not solely due to LingoDeer. Control and intervention classrooms were planned, but no section or condition key is included in the participant-level analytic file except for grade. The intended comparison cannot be used in a pooled paired analysis, and may be due to maturation, normal instruction, repeated testing, teaching to the test, attendance selection, regression to the mean, and unequal exposure. A causal or between-group claim must be accompanied by recovery and verification of the condition key as well as actual-dose information.

Implications for practice. The instructional protocol provides a starting plan that is replicable: six units of curriculum aligned instruction, teacher introduction, communicative follow-up, independent practice for 15-20 minutes daily, and weekly review. Future implementations need to record actual sessions, attendance, time on task, module completion, teacher fidelity,

home access and technical interruptions. Before making decisions about adopting the plan, it should be evaluated using parallel assessment forms, a delayed retention measure, and a prespecified control comparison.

Limitations and directions for research. First, the class-section and condition identifiers were not retained in the participant-level analytic file and thus were not available for use of the planned controls. Second, the actual dose, adherence, usage logs, fidelity, and concurrent instruction were not measured or retained. Third, the same test constructed locally was administered twice, there was no delayed test to measure retention, and the post test had ceiling compression and alpha = -.395. Fourth, the enrollment summary shows four students with SEN, and the analytic file does not contain an individual SEN field, and no exclusion was applied for SEN. Fifth, the contextual questionnaire included 71 respondents, whereas the matched achievement sample included 70 because one questionnaire respondent was absent from the achievement posttest and could not be included in the paired analysis.

CONCLUSIONES

Three conclusions follow from the analysis of the primary participant-level data.

First, the baseline diagnosis establishes a contextual need for additional support. Across digital-mediation items, 45.1%-67.6% of 71 respondents selected 'Never'; across self-perceived lexical-performance items, 56.3%-63.4% selected 'Never'. Audio was the most frequently preferred resource (35/71, 49.3%). These questionnaire findings answer the baseline research question and should not be presented as intervention outcomes.

Second, the primary score data support a large pooled descriptive change: pretest M = 12.26 (SD = 5.12), posttest M = 25.80 (SD = 1.42), mean difference = 13.54 (SD = 5.22), $t(69) = 21.69$, $p < .001$, with AAR/DAR increasing from 7.1% to 100.0%. The corresponding $dz = 2.59$ is a within-person statistic and must not be presented as a controlled application effect.

Third, primary-data status does not resolve the design and measurement limitations. Condition identifiers and verified actual dose are missing; the same test was repeated; final scores are near the ceiling; and posttest alpha is negative. The defensible conclusion is that vocabulary scores increased during the instructional period. A causal claim about LingoDeer requires a verified control/experimental key and a prespecified controlled analysis or a new controlled replication.

AGRADECIMIENTOS (OPCIONAL)

The authors thank the management team, teaching staff, students, and families of the Guardiana de los Saberes San Gabriel de Piquiucho Educational Unit for their participation and for authorizing the implementation.

REFERENCIAS BIBLIOGRÁFICAS

1. Nation ISP. Learning vocabulary in another language. 2nd ed. Cambridge: Cambridge University Press; 2013. https://doi.org/10.1017/CBO9781139858656
2. Schmitt N. Instructed second language vocabulary learning. Language Teaching Research. 2008;12(3):329-63. https://doi.org/10.1177/1362168808089921
3. Schmitt N, Schmitt D. Vocabulary in language teaching. 2nd ed. Cambridge: Cambridge University Press; 2020. https://doi.org/10.1017/9781108569057
4. Webb S, Nation P. How vocabulary is learned. Oxford: Oxford University Press; 2017.
5. Krashen SD. The input hypothesis: issues and implications. London: Longman; 1985.
6. Vygotsky LS. Mind in society: the development of higher psychological processes. Cambridge (MA): Harvard University Press; 1978.
7. Burston J, Giannakou K. MALL language learning outcomes: a comprehensive meta-analysis 1994-2019. ReCALL. 2022;34(2):147-68. https://doi.org/10.1017/S0958344021000240
8. Hao T, Wang Z, Ardasheva Y. Technology-assisted vocabulary learning for EFL learners: a meta-analysis. Journal of Research on Educational Effectiveness. 2021;14(3):645-67. https://doi.org/10.1080/19345747.2021.1917028
9. Lin JJ, Lin H. Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis. Computer Assisted Language Learning. 2019;32(8):878-919. https://doi.org/10.1080/09588221.2018.1541359
10. Mahdi HS. Effectiveness of mobile devices on vocabulary learning: a meta-analysis. Journal of Educational Computing Research. 2018;56(1):134-54. https://doi.org/10.1177/0735633117698826
11. Peng H, Jager S, Lowie W. Narrative review and meta-analysis of MALL research on L2 skills. ReCALL. 2021;33(3):278-95. https://doi.org/10.1017/S0958344020000221
12. Yu A, Trainin G. A meta-analysis examining technology-assisted L2 vocabulary learning. ReCALL. 2022;34(2):235-52. https://doi.org/10.1017/S0958344021000239
13. Zhou Y, Zhou M. A meta-analysis on mobile-assisted vocabulary learning: do mobile applications help? ReCALL. 2026;38(1):75-93. https://doi.org/10.1017/S0958344025100335
14. Kukulska-Hulme A, Shield L. An overview of mobile assisted language learning: from content delivery to supported collaboration and interaction. ReCALL. 2008;20(3):271-89. https://doi.org/10.1017/S0958344008000335
15. Loewen S, Crowther D, Isbell DR, Kim KM, Maloney J, Miller ZF, et al. Mobile-assisted language learning: a Duolingo case study. ReCALL. 2019;31(3):293-311. https://doi.org/10.1017/S0958344019000065
16. Burston J, Athanasiou A, Giannakou K. Quantitative experimental L2 acquisition MALL studies: a critical evaluation of research quality. ReCALL. 2024;36(1):22-39. https://doi.org/10.1017/S0958344023000149
17. Hou Z, Aryadoust V. A review of the methodological quality of quantitative mobile-assisted language learning research. System. 2021;100:102568. https://doi.org/10.1016/j.system.2021.102568
18. Burston J. Twenty years of MALL project implementation: a meta-analysis of learning outcomes. ReCALL. 2015;27(1):4-20. https://doi.org/10.1017/S0958344014000159

19. Nakata T. Computer-assisted second language vocabulary learning in a paired-associate paradigm: a critical investigation of flashcard software. Computer Assisted Language Learning. 2011;24(1):17-38. https://doi.org/10.1080/09588221.2010.520675
20. Stockwell G, Hubbard P. Some emerging principles for mobile-assisted language learning. Monterey (CA): The International Research Foundation for English Language Education; 2013.
21. LingoDeer. LingoDeer: learn languages [Internet]. 2023 [cited 2026 Jul 23]. Available from: https://www.lingodeer.com/
22. Triwiyanti I, Dalimunte AA. Learning English vocabulary by LingoDeer: case study of Indonesian junior high school students. Acuity: Journal of English Language Pedagogy, Literature and Culture. 2025;10(3):276-89. https://doi.org/10.35974/acuity.v10i3.4166
23. Ministerio de Educación del Ecuador. English as a foreign language curriculum [Internet]. Quito: Ministerio de Educación; 2016 [cited 2026 Jul 23]. Available from: https://educacion.gob.ec/wp-content/uploads/downloads/2016/08/0-EFL_v2.pdf
24. Council of Europe. Common European Framework of Reference for Languages: learning, teaching, assessment, companion volume. Strasbourg: Council of Europe Publishing; 2020.
25. Mishra P, Koehler MJ. Technological pedagogical content knowledge: a framework for teacher knowledge. Teachers College Record. 2006;108(6):1017-54. https://doi.org/10.1111/j.1467-9620.2006.00684.x
26. Koehler MJ, Mishra P. What is technological pedagogical content knowledge? Contemporary Issues in Technology and Teacher Education. 2009;9(1):60-70.
27. Li Y, Hafner CA. Mobile-assisted vocabulary learning: investigating receptive and productive vocabulary knowledge of Chinese EFL learners. ReCALL. 2022;34(1):66-80. https://doi.org/10.1017/S0958344021000161