

Augmented Reality-Supported Experiential Learning in Primary EFL: Effects on Vocabulary Achievement, Motivation, and Learner Experience

Mohammed Z M Qaddoura¹, Mohd Fadzil Abdul Hanid¹, Mohd Nihra Haruzuan Mohamad Said¹, Huan Yik Lee²

¹Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia, 81300, Skudai, Johor, Malaysia

²Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, 81300, Skudai, Johor, Malaysia

Correspondence: Mohd Fadzil Abdul Hanid, Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia, 81300, Skudai, Johor, Malaysia.

Received: September 2, 2026

Accepted: October 17, 2026

Online Published: October 22, 2026

doi:10.11114/jets.v15i1.9265

URL: <https://doi.org/10.11114/jets.v15i1.9265>

Abstract

Augmented Reality (AR) can enhance English as a Foreign Language (EFL) learning, yet evidence on primary learners in multicultural contexts remains limited. This study investigated the effects of AR-supported experiential learning on vocabulary achievement, motivation, and learner perceptions among 120 grade six EFL learners in Arab schools in Malaysia. An explanatory sequential mixed-methods quasi-experimental design was used, with 62 learners receiving AR-supported instruction structured around Kolb's experiential learning cycle and 58 receiving conventional instruction. Vocabulary achievement was measured using pre- and post-tests, while motivation was assessed through the ARCS dimensions of Attention, Relevance, Confidence, and Satisfaction. Quantitative data were analysed using Analysis of Covariance (ANCOVA) and Multivariate Analysis of Covariance (MANCOVA), followed by semi-structured interviews with 10 purposively selected learners. Both groups improved their vocabulary achievement, with a statistically significant adjusted advantage for the AR group, $F(1, 117) = 75.68, p < .001$, partial $\eta^2 = .393$. The AR condition produced a significant multivariate effect on motivation, with particularly strong effects for Attention and Satisfaction. Qualitative findings indicated that AR increased engagement, enjoyment, perceived retention, autonomy, and willingness to participate, although learners experienced technical difficulties and an initial learning curve. Overall, AR-supported experiential learning was associated with substantially higher vocabulary achievement and stronger motivation than conventional instruction, while enhancing learners' engagement and perceived learning experiences.

Keywords: augmented reality, experiential learning, vocabulary acquisition, motivation, English as a foreign language

1. Introduction

The rapid pace of technological development has made adopting digital tools an urgent concern for education systems (Haleem et al., 2022). Among these tools, Augmented Reality (AR), which superimposes digital information onto the physical environment, has emerged as a particularly promising instructional medium (Chang et al., 2022). Its appeal is clearest in language education, where English proficiency has become a prerequisite for international participation. However, conventional teaching methods continue to struggle to motivate learners (Larsen-Freeman & Freeman, 2008).

These difficulties are acute for grade six learners of English as a Foreign Language (EFL) in Arabian schools in Malaysia, where three problems recur: the structural distance between Arabic and English, restricted opportunity to practise English outside school, and low motivation arising from monotonous traditional instruction (Azar & Tanggaraju, 2020; Daud et al., 2020; Jaliyya & Idrus, 2017; Rass, 2024). Remedies proposed to date, such as culturally relevant materials and enhanced teacher preparation (Aronson & Laughter, 2015), have not meaningfully improved participation or satisfaction. Note that both are necessary conditions for successful language learning, and limited training in contemporary pedagogies continues to depress attainment and enthusiasm alike (Elena V. Frolova et al., 2020).

Nevertheless, the educational impact of AR at primary level remains largely untested, since the existing evidence base is concentrated in secondary and post-secondary settings (Akçayır & Akçayır, 2017). Little is known about how AR shapes vocabulary learning, engagement, and learner perceptions in Malaysian Arabian schools, which operate under distinctive cultural and linguistic conditions (Saud et al., 2023), and limited empirical evidence is available regarding the systematic integration of AR with experiential learning in primary EFL settings (Belda-Medina & Calvo-Ferrer, 2022). Therefore, this study examines how AR-based experiential learning affects English vocabulary acquisition

among grade six EFL learners in Malaysian Arabian schools. Three questions are addressed. First, what is the effect of AR-based experiential learning on vocabulary achievement? Second, how does it influence motivation as measured by the ARCS model of Attention, Relevance, Confidence, and Satisfaction? Third, what attitudes do learners hold towards AR in the EFL classroom, and how do they perceive its effectiveness and their own satisfaction?

2. Literature Review

2.1 *Augmented Reality as an Educational Resource*

AR creates interactive learning environments by projecting digital materials onto physical space (Yuen et al., 2011). Empirical work indicates that this capacity enhances engagement, knowledge retention, and motivation by providing authentic language contexts and personalised feedback (AlGerafi et al., 2023; Godwin-Jones, 2004). Given that learners can rehearse vocabulary in situations that resemble real use while reflecting on their own learning, AR supports acquisition through practical experience rather than passive exposure. Furthermore, studies in second language education report consistent gains in engagement and acquisition (Chan, 2022; Hung et al., 2016; Wen, 2020; Y. Zhang & Miao, 2025). A recent systematic review further corroborates these findings, reporting positive outcomes in motivation, engagement, and proficiency while noting that primary-level evidence remains scarce (Qaddoura et al., 2026).

This work particularly focuses on vocabulary. Reviews of AR-assisted vocabulary learning report generally positive effects on recall, attributing them to the combination of visual and verbal encoding and to the embodied quality of manipulating virtual referents (Ma et al., 2025; M. Zhang et al., 2025). Game-based AR applications have likewise been associated with gains in motivation among elementary EFL learners in both rural and urban settings (Liao et al., 2023). Teachers, however, report that realising these benefits depends on marker reliability, device availability, and classroom management (Salehi, 2025), and evidence remains stronger for motivational than for achievement outcomes.

2.2 *Experiential Learning and Constructivist Foundations*

Experiential learning, which privileges practical activity and connection to authentic situations (Chan, 2022), provides a theoretical basis for leveraging these affordances and enables AR to create interactive vocabulary-learning spaces that conventional instruction cannot provide (Hung et al., 2016). Kolb's (2014) cycle of concrete experience, reflective observation, abstract conceptualisation and active experimentation supplies the structure through which such affordances become pedagogy rather than novelty. This orientation is consistent with constructivist accounts of language learning, in which knowledge is built through learner activity and negotiated meaning rather than transmitted (Bamanger, 2025). By contrast, existing digital resources remain largely multimedia and lack the interactive, experiential components that immersive classrooms require (Nattida, 2021).

2.3 *Theoretical Framework*

The study is framed by two complementary models (Kolb, 2014). Experiential learning theory specifies how learning proceeds, describing a cycle in which concrete experience is transformed into knowledge through reflection, conceptualisation, and active testing. Thus, it prescribes what learners should do with an augmented object rather than merely how that object should appear (Keller, 2009). ARCS model specifies the conditions under which learners will invest effort, treating motivation as the product of attention gained, relevance perceived, confidence built, and satisfaction experienced. The two models address different questions, and combining them allows us to examine AR's affordances as both pedagogy and motivational design. In the present study, Kolb's cycle governed the structure of the classroom activities, the ARCS dimensions provided the outcome framework against which learner motivation was measured, and the qualitative phase examined how learners themselves connected the two.

2.4 *Barriers to Implementation and the Research Gap*

Implementation nonetheless remains uneven. In Malaysian Arabic schools, limited access to suitable devices, insufficient teacher training, and unresolved difficulties in integrating AR into the prescribed curriculum constrain adoption (Ghani et al., 2025; Park & Sung, 2013). Moreover, few studies have examined how AR can advance experiential learning for English language learners in culturally diverse settings or simultaneously traced its effects on vocabulary growth, engagement, perceived effectiveness, and satisfaction.

Although previous studies have generally reported positive effects of AR on vocabulary learning and motivation, the empirical evidence remains mixed and context-dependent. For instance, Liao et al. (2023) reported favourable learning and motivational outcomes in AR-supported language activities. In contrast, Binhomran & Altalhab (2021) discovered that, despite higher mean vocabulary acquisition and retention scores among primary EFL learners using AR, these differences were not statistically significant compared with conventional instruction. Similarly, Jia et al. (2025) determined significant immediate vocabulary gains but no significant between-group differences in delayed retention, raising questions about the consistency and sustainability of AR-related learning benefits. Chang et al. (2022) further highlighted variations in AR effectiveness associated with intervention characteristics, particularly instructional duration and design. Collectively, these

findings suggested that technological immersion alone does not guarantee superior learning achievement and that pedagogical integration requires closer examination. Nevertheless, few empirical studies have systematically integrated AR with Kolb's four-stage experiential learning cycle while examining vocabulary achievement, ARCS-based motivation, and learners' experiences among primary EFL learners in multicultural Arab schools in Malaysia. Note that addressing this gap requires a pedagogically grounded mixed-methods investigation that evaluates both measurable learning outcomes and learners' experiences of AR-supported experiential instruction. The present study addresses this gap by evaluating an AR-based experiential learning intervention in an under-researched primary EFL population.

3. Methodology

3.1 Research Design

The three research questions require different types of inference. Questions 1 and 2 examine cause-and-effect relationships between instruction and tangible outcomes (vocabulary achievement and motivation). In contrast, Question 3 concerns attitudes, perceptions, and satisfaction, which Likert-scale measures alone cannot adequately explain. Therefore, an explanatory sequential mixed-methods design was adopted (Creswell, 2018), in which a quantitative phase preceded and framed a qualitative phase.

The quantitative phase used a quasi-experimental, non-equivalent control-group pre-test–post-test design with existing Grade Six classes rather than individual random assignment. Both instructional conditions were implemented within intact classes drawn from the three participating schools. Vocabulary achievement was assessed using pre- and post-tests, and motivation using a questionnaire based on Keller's ARCS model. The qualitative phase comprised semi-structured interviews conducted after the intervention with a purposive sample of 10 students from the experimental condition. Consequently, quantitative and qualitative findings were interpreted to contextualise differences in adjusted vocabulary and motivation.

3.2 Participants and Sampling

The target population comprised grade six learners in Arab schools in Malaysia, aged 11 to 12, learning EFL within a distinctive socio-cultural and linguistic environment. The accessible population consisted of Grade Six EFL learners from three private Arab schools in Kuala Lumpur and Selangor that agreed to participate and possessed the necessary tablets.

The study involved 120 Grade Six EFL learners from three private Arab schools in Kuala Lumpur and Selangor. Participants were organised into six existing intact classes, with two classes recruited from each school. Within each school, one intact class received AR-supported experiential instruction, while the other received conventional vocabulary instruction. School A contributed 21 students to the experimental group and 19 students to the control group. School B contributed 20 students to each group, while School C contributed 21 students to the experimental group and 19 to the control group. Overall, the experimental group comprised 62 students across three intact classes, while the control group comprised 58 students across three intact classes.

Three English language teachers were involved, with one teacher from each school delivering instruction to both the experimental and control classes within their respective schools. This arrangement was intended to reduce differences between teachers within each school while preserving the existing classroom structure. Group allocation was non-random, as students remained in their established classes throughout the intervention. Nevertheless, potential teacher expectancy, cross-condition contamination, classroom characteristics, and school-level influences could not be eliminated. A previously reported power calculation referred to an independent-samples comparison with a medium effect size ($d = 0.50$), a two-sided $\alpha = .05$, and power of $1 - \beta = .80$ (Mohanasundaram et al., 2024). Under these assumptions, a conventional two-group power calculation gives approximately 128 students in total (64 per group), rather than 102. The realised sample of 120 falls below this illustrative individual-level target, and the calculation does not account for the additional uncertainty associated with intact classroom clusters.

For the qualitative phase, 10 learners were selected by maximum-variation purposive sampling in order to capture a range of experiences rather than statistical representativeness: four high achievers scoring above 85% at post-test, three medium achievers scoring between 65% and 84%, and three low achievers scoring below 65%, with five male and five female participants. Note that written parental consent and each learner's own assent were obtained before participation.

3.3 Instructional Conditions

3.3.1 AR-Supported Experiential Instruction

The intervention employed a custom marker-based application, AR Vocabulary Explorer, developed in Unity with the Vuforia software development kit and pre-installed on Android tablets. Each target word was linked to a unique printed marker embedded in a flashcard. Correspondingly, scanning the marker generated an interactive 3D representation that learners could rotate, zoom, and manipulate, while audio pronunciation and contextual sentences reinforced connections

among visual form, spoken form, and meaning. All content was pre-loaded, allowing offline operation.

Activities followed the experiential learning cycle (Kolb, 2014). During concrete experience, learners scanned markers and interacted with three-dimensional objects in their physical environment. In reflective observation, they described the objects in pairs and related them to prior knowledge. During abstract conceptualisation, they connected the representations to meanings, spellings, pronunciations, and contextual uses. Meanwhile, during active experimentation, they applied the new words in sentence construction, scenario description, and role-play. The intervention ran for six weeks, with three 45-minute sessions each week, for a total of 18 sessions. Note that each session introduced five to seven items, and earlier words were revisited through cumulative review.

3.3.2 Conventional Instruction (Control Condition)

The control group received conventional teacher-led vocabulary instruction without the AR Vocabulary Explorer application or interactive marker-based three-dimensional objects. The instruction took place over six weeks, with three 45-minute sessions per week, for a total of 18 sessions. Each session covered the same five to seven target vocabulary items as the experimental group to maintain consistency in learning content. Conventional teaching materials included English textbooks, printed picture flashcards, whiteboards, and worksheets. Each lesson began with a brief review of previously learned vocabulary, followed by teacher-led explanations of new words, pronunciation modelling, and oral repetition. Consequently, students completed guided vocabulary activities, including word–picture matching, sentence completion, and sentence writing. Teachers provided verbal corrections and written feedback to reinforce vocabulary understanding and accurate usage. Both groups followed the same instructional schedule, covered the same vocabulary content, and completed identical pre- and post-test assessments. However, unlike the experimental group, which engaged in interactive three-dimensional vocabulary exploration structured around Kolb's experiential learning cycle, the control group learned through conventional teacher-guided explanations and textbook-based practice. This arrangement was intended to facilitate a comparison between AR-supported experiential learning and conventional vocabulary instruction while maintaining comparable instructional duration and learning content.

3.4 Instruments and Quantitative Analysis

Vocabulary achievement was measured by parallel pre- and post-tests scored out of 100. Motivation was measured with a 20-item questionnaire, five items per subscale, based on Keller's ARCS model and using a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Its reliability is reported in Table 5.

Before inferential analysis, normality was assessed using the Shapiro–Wilk test and Q–Q plots, and homogeneity of variance was assessed using Levene's test. Linearity and homogeneity of regression slopes were examined for ANCOVA. As Table 1 illustrates, the experimental group pre-test did not meet normality ($W = 0.963$, $p = .041$), whereas the other reported normality, variance, and regression-slope diagnostics were acceptable. However, intact-class allocation implies that observations within the same class may be correlated. The reported individual-level assumption tests do not establish independence between students.

Descriptive statistics were calculated for both groups, and an independent-samples t-test assessed the baseline group difference. The primary analysis was an ANCOVA with post-test score as the dependent variable, instructional group as the fixed factor, and pre-test score as the covariate. Motivation was analysed by MANCOVA with the four ARCS subscales as dependent variables and pre-test motivation scores as covariates, followed by univariate ANCOVAs for each subscale. Effect sizes were reported as partial eta squared, interpreted as small (.01), medium (.06), and large (.14), and as Cohen's d . Paired-samples t-tests examined within-group change as complementary evidence rather than as a substitute for the between-group comparison.

The reported ANCOVA and MANCOVA were conducted at the student level and did not explicitly model classroom or school clustering. Consequently, standard errors, confidence intervals, and statistical significance may be sensitive to within-class dependence. Given the small number of classes and schools, reliable cluster-adjusted estimates cannot be assumed without further analysis. The results are therefore interpreted with caution as pre-test-adjusted associations with instructional conditions, rather than as definitive causal estimates of AR's isolated effect.

3.5 Qualitative Analysis

Interviews lasted 25–35 minutes and were conducted in Arabic, the participants' first language, to enable them to express their beliefs and experiences freely. Transcripts were back-translated into English to preserve meaning and linguistic accuracy. The transcripts were analysed thematically following the six phases described by Braun & Clarke (2006): familiarisation with the data, line-by-line coding, generation of themes, review of themes against the data extracts, definition and naming of themes, and reporting. Two raters independently coded 30% of the interviews. Initial agreement was 87.3%, and Cohen's kappa was 0.84 ($p < .001$), which Landis & Koch (1977) classified as almost perfect. Discrepancies were resolved through discussion and by making the code definitions operational. Trustworthiness was

further supported by conducting the interviews in the participants' first language, retaining verbatim extracts as evidence for each theme, and using maximum-variation sampling to ensure both lower- and higher-achieving learners were represented. Participants were identified by code throughout, and no identifying information was retained in the transcripts.

Table 1. Assumption Testing Results

Assumption	Test	Experimental Group	Control Group	Decision
Normality (pre-test)	Shapiro-Wilk	W = .963, $p = .041$	W = .968, $p = .092$	Violated (experimental) / Assumed (control)
Normality (post-test)	Shapiro-Wilk	W = .981, $p = .124$	W = .975, $p = .108$	Assumed
Homogeneity of variances (pre-test)	Levene	F(1, 118) = 1.338, $p = .250$	-	Assumed
Homogeneity of variances (post-test)	Levene	F(1, 118) = 0.876, $p = .351$	-	Assumed
Homogeneity of regression slopes	Interaction term	F(1, 116) = 0.419, $p = .519$	-	Met

4. Results

4.1 Vocabulary Achievement

Table 2. Descriptive Statistics for Vocabulary Achievement

Group	N	Pre-test M (SD)	Post-test M (SD)	Mean Gain (Post-Pre)
Experimental (AR)	62	52.95 (10.88)	78.23 (9.56)	25.28
Control (Traditional)	58	51.45 (9.82)	62.43 (10.12)	10.98

Note: Maximum possible score = 100.

The two groups began the study with comparable vocabulary knowledge (Table 2). An independent-samples t -test observed no statistically significant baseline difference, $t(118) = 0.791$, $p = .431$, Cohen's $d = 0.145$. At post-test, the experimental group had a mean of 78.23 ($SD = 9.56$), compared with 62.43 ($SD = 10.12$) for the control group, a gain of 25.28 points versus 10.98 points.

After controlling for pre-test vocabulary scores, the ANCOVA indicated a statistically significant between-group difference in post-test achievement, $F(1, 117) = 75.68$, $p < .001$, partial $\eta^2 = .393$. The AR-supported group achieved higher adjusted post-test scores than the control group, and instructional condition accounted for approximately 39.3% of the residual-adjusted outcome variance, representing a large partial effect under conventional benchmarks (Cohen, 1988).

Table 3. Within-Group Improvements

Group	t-value	df	p-value	Cohen's d	Interpretation
Experimental (AR)	14.282	61	$< .001$	2.56	Large
Control	6.347	57	$< .001$	1.11	Large

Both groups improved significantly from pre-test to post-test (Table 3). The experimental group yielded $t(61) = 14.282$, $p < .001$, $d = 2.56$, while the control group yielded $t(57) = 6.347$, $p < .001$, $d = 1.11$. Although both effects fall within the conventionally large range, the within-group effect obtained under the AR condition was more than twice that obtained under traditional instruction.

Table 4. Adjusted Post-test Means (ANCOVA)

Group	Adjusted M	SE	95% CI
Experimental (AR)	78.16	1.25	[75.69, 80.64]
Control	62.50	1.29	[59.94, 65.06]

Table 4 reports post-test means adjusted for pre-test vocabulary scores. The experimental group recorded an adjusted mean of 78.16 ($SE = 1.25$, 95% CI [75.69, 80.64]), compared with 62.50 ($SE = 1.29$, 95% CI [59.94, 65.06]) in the control group. The adjusted mean difference was 15.66 points (95% CI [12.09, 19.22]), favouring AR-supported experiential learning.

4.2 Motivation

Table 5. Instrument Reliability

Subscale	Cronbach's α (pre-test)	Cronbach's α (post-test)	McDonald's ω
Attention	0.841	0.879	0.885
Relevance	0.823	0.862	0.868
Confidence	0.799	0.841	0.849
Satisfaction	0.814	0.875	0.879
Total	0.912	0.941	0.945

As Table 5 illustrates, Cronbach's *alpha* ranged from 0.799 to 0.841 at pre-test and from 0.841 to 0.879 at post-test across the four subscales, with McDonald's *omega* ranging from 0.849 to 0.885. The total scale returned an alpha of 0.912 at pre-test and 0.941 at post-test, with an omega of 0.945. Therefore, all coefficients exceeded the conventional threshold of 0.70, and the total scale may be regarded as demonstrating excellent internal consistency.

Table 6. Descriptive Statistics for Motivation Scores (Post-intervention)

ARCS Subscale	Experimental Group (n=62) <i>M</i> (<i>SD</i>)	Control Group (n=58) <i>M</i> (<i>SD</i>)
Attention	4.40 (0.55)	3.21 (0.72)
Relevance	4.26 (0.60)	3.45 (0.68)
Confidence	4.03 (0.68)	3.38 (0.74)
Satisfaction	4.48 (0.51)	3.29 (0.69)
Total Motivation	4.29 (0.49)	3.33 (0.61)

Note: Scale ranges from 1 (low) to 5 (high).

Post-intervention scores favoured the experimental group on every ARCS subscale (Table 6), with a total motivation mean of 4.29 (*SD* = 0.49) against 3.33 (*SD* = 0.61) for the control group. The widest margin was in Satisfaction, and the narrowest in Confidence.

Table 7. MANCOVA Results

Statistical Test	Value	F	df1	df2	p	Partial η^2
Wilks' Λ	0.472	26.119	4	112	< .001	0.483
Pillai's Trace	0.528	26.119	4	112	< .001	0.483
Hotelling's Trace	1.119	26.119	4	112	< .001	0.483
Roy's Largest Root	1.119	26.119	4	112	< .001	0.483

With pre-test motivation scores entered as covariates, the MANCOVA produced a significant multivariate effect, $F(4, 112) = 26.119$, $p < .001$, partial $\eta^2 = 0.483$, and all four multivariate criteria converged on the same outcome (Table 7). Hence, instructional condition accounted for approximately 48.3% of the variance in the combined motivation measures.

Table 8. Univariate ANCOVAs

Subscale	$F(1, 115)$	p	Partial η^2	Effect Size
Attention	61.234	< .001	0.347	Large
Relevance	34.876	< .001	0.233	Large
Confidence	17.921	< .001	0.135	Medium-large
Satisfaction	68.452	< .001	0.373	Large

The follow-up univariate analyses confirmed that the effect held for each subscale considered separately (Table 8): Attention, $F(1, 115) = 61.234$, $p < .001$, partial $\eta^2 = 0.347$; Relevance, $F(1, 115) = 34.876$, $p < .001$, partial $\eta^2 = 0.233$; Confidence, $F(1, 115) = 17.921$, $p < .001$, partial $\eta^2 = 0.135$; and Satisfaction, $F(1, 115) = 68.452$, $p < .001$, partial $\eta^2 = 0.373$.

Table 9. Magnitude of Differences (Cohen's *d*)

Subscale	Cohen's <i>d</i>	95% CI for <i>d</i>	Interpretation
Attention	1.85	[1.42, 2.28]	Very large
Relevance	1.24	[0.85, 1.63]	Large
Confidence	0.89	[0.51, 1.27]	Large
Satisfaction	1.98	[1.54, 2.42]	Very large
Total	1.71	[1.29, 2.13]	Very large

Standardised effect sizes (Table 9) ranged from $d = 0.89$ (95% CI [0.51, 1.27]) for Confidence, through $d = 1.24$ (95% CI [0.85, 1.63]) for Relevance, to $d = 1.85$ (95% CI [1.42, 2.28]) for Attention and $d = 1.98$ (95% CI [1.54, 2.42]) for Satisfaction. Total motivation yielded $d = 1.71$ (95% CI [1.29, 2.13]). Given that none of these intervals includes zero, the AR condition consistently indicated an advantage across all four dimensions.

4.3 Learner Attitudes and Perceptions

Thematic analysis of the 10 interviews produced five superordinate themes, each with two subthemes, which are summarised in Table 10 with representative extracts.

Table 10. Thematic Analysis Summary

Superordinate Theme	Subthemes	Frequency (n=10)	Representative Quote
1. Experiential Engagement	1.1 Interactive immersion	10/10	"The AR made it feel like the words came to life... I can touch them, rotate them. In class, we just copy." (S03, Male, High achiever)
	1.2 Reduced affective filter	9/10	"I felt brave enough to commit errors because the AR app gave me the chance to do it without making any noise." (S07, Female, Low achiever)
2. Perceived Effectiveness	2.1 Vocabulary retention	10/10	"After that, whenever I look at the word 'mountain,' I remember the 3D mountain model I created in AR." (S01, Male, Medium achiever)
	2.2 Contextual understanding	8/10	"In the past, I learned vocabulary by heart. Now, I know the word's context." (S05, Female, High achiever)
3. Motivational Drivers	3.1 Novelty and curiosity	8/10	"I always looked forward to each day... which AR word will we use today?" (S02, Female, Medium achiever)
	3.2 Autonomy and control	7/10	"It is good that we could study at our own pace without anyone else watching us." (S09, Male, Low achiever)
4. Satisfaction Factors	4.1 Enjoyment as learning catalyst	9/10	"It turned out that learning could be fun. I advised my brother to visit my English class." (S06, Female, High achiever)
	4.2 Technological appreciation	10/10	"This AR experience was unique... we did something else than usual when visiting other classes." (S04, Male, Medium achiever)
5. Challenges and Limitations	5.1 Technical difficulties	4/10	"I think there were moments where it was not easy for the app to work. Twice during AR task, my tablet got frozen." (S08, Female, Low achiever)
	5.2 Initial learning curve	5/10	"At first, I had problems scanning the marker. But then, it was very simple." (S10, Male, Medium achiever)

All ten participants reported experiential engagement and described vocabulary learning as becoming participatory rather than receptive. The shift was felt most strongly by lower-achieving learners, for whom AR reduced classroom anxiety: "In a normal classroom, teachers would ask questions, and I would cover my face because I did not know. But with AR, I could just play around. The words stayed because I was doing something." (S07, female, low achiever)

Perceived effectiveness was attributed to the visual and spatial character of the medium, which several learners described in embodied terms: "When we learned 'enormous,' the AR made the elephant gigantic on my desk. I felt tiny. I remember because I use my body and not just my mind." (S03, male, high achiever)

On the other hand, participants distinguished motivational drivers from mere novelty. Participants separated their initial excitement from the sustained engagement that followed it: "After one week, everything was different. It was not just 'wow' technology anymore but 'wow, I can really learn English.'" (S05, female, high achiever)

Satisfaction was closely tied to perceived progress and was expressed most strongly by learners who had previously struggled: "My score on the quiz increased from 45 to 82. Seeing that, I just could not believe it. AR made me prove to myself that I am not stupid when it comes to English." (S09, male, low achiever)

Correspondingly, a minority reported challenges. Four participants cited technical difficulties with tablet responsiveness and marker recognition, and five referred to an initial learning curve, but none regarded these as outweighing the benefits: "Yes, it was slow sometimes ... but even though there were issues, it was still better than the textbook." (S08, female, low achiever)

Table 11. Qualitative–Quantitative Integration (Meta-inference)

Quantitative Finding	Qualitative Explanation	Convergence
Large effect on Satisfaction ($\eta^2 = 0.373$)	"Learning was actually fun" + "AR helped me prove to myself"	Confirmed
Large effect on Attention ($\eta^2 = 0.347$)	"Words were coming alive" + "Every day I was excited"	Confirmed
Smaller effect on Confidence ($\eta^2 = 0.135$)	"Initial learning curve" + "Not scared to make mistakes"	Partial (confidence improved but required adaptation period)
Significant between-group vocabulary advantage (partial $\eta^2 = .393$)	Reported visual-spatial recall and contextual understanding are consistent with higher adjusted AR-group vocabulary scores; interviews do not by themselves establish causation	Convergent (qualitative accounts support the direction of the quantitative difference)

Table 11 sets the qualitative themes against the quantitative findings. Interviews helped contextualise the larger effects on Satisfaction and Attention and the relatively smaller effect on Confidence, including the initial period of adaptation described by several learners. Learners' accounts of visual recall and contextual understanding also aligned with the statistically significant adjusted between-group vocabulary difference. However, interviews alone cannot establish the causal mechanism underlying that difference.

5. Discussion

This study examined the effects of AR-supported experiential learning on the vocabulary gains and motivation of primary EFL learners. Learners' descriptions of visual-spatial recall and contextual word use were also consistent with the statistically significant adjusted vocabulary advantage observed in the AR group, although interview accounts cannot establish a causal mechanism.

5.1 Vocabulary Achievement

The adjusted ANCOVA comparison indicates a statistically significant advantage for AR-supported experiential learning in vocabulary achievement, $F(1, 117) = 75.68$, $p < .001$, partial $\eta^2 = .393$. An adjusted mean difference of 15.66 points on the 100-point vocabulary test represents a substantial educational contrast. At the same time, the large partial effect suggests that group membership remained associated with achievement after adjustment for pre-test scores. Nevertheless, statistical significance does not remove limitations arising from intact-class assignment, a six-week intervention, and an immediate post-test. More generally, the distinction between absence of evidence and evidence of absence (Altman & Bland, 1995) underscores the need to interpret inferential results accurately and cautiously.

The six-week exposure was sufficient to yield an adjusted between-group difference. However, the durability of that difference remains unknown. Parmaxi & Demetriou (2020) highlighted how task characteristics, exposure time, and novelty may shape AR effectiveness in language learning. Therefore, delayed post-tests and longer implementation are needed to determine whether vocabulary advantages persist. Evidence from quasi-experimental intervention research also demonstrates variation in observed intervention effects (Doyle et al., 2022), reinforcing the need for replication rather than assuming that a short-term achievement advantage generalises across settings.

The conventional instruction group also improved substantially, with a within-group effect size of $d = 1.11$. This indicates that the comparison condition was active and educationally useful. Crucially, the adjusted between-group contrast nevertheless favoured the AR group, suggesting an added vocabulary benefit beyond that observed with conventional teaching. The interpretation therefore concerns the comparative value of two instructional approaches rather than the failure of the traditional condition (Doughty & Long, 2003).

The instructional contrast also combined differences in delivery technology and pedagogy: the AR condition incorporated object manipulation and a sequenced experiential cycle, whereas the comparator followed routine instruction. Parmaxi & Demetriou (2020) highlighted how task characteristics and exposure can shape augmented-reality learning. Doyle et al. (2022) documented differences between quasi-experimental intervention outcomes. Accordingly, the adjusted group difference should not be interpreted as the AR display's isolated contribution.

Note that the strength of the comparison further depends on what occurred during routine lessons. Although the control group improved, its gains do not establish equivalent teacher contact, practise opportunities, or feedback. These aspects require verification from the original classroom records, especially because instructional practices can shape language-learning outcomes (Doughty & Long, 2003).

Taken together, the reported student-level ANCOVA supports a short-term adjusted association between AR-supported experiential learning and higher vocabulary achievement relative to conventional instruction. Nevertheless, non-random intact-class assignment, possible teacher and school influences, and unmodelled classroom clustering preclude an unqualified causal attribution to AR alone. The intervention combined a technology component with a structured experiential pedagogy. This study cannot isolate the separate contributions of each.

5.2 Motivation

The motivational findings were unambiguous. The multivariate effect (Wilks' $\Lambda = 0.472$, partial $\eta^2 = 0.483$) indicates that instructional condition explained 48.3% of the multivariate variance in motivation, and all four subscales returned significant univariate effects.

The effect on Attention ($d = 1.85$) aligns with established evidence that AR captures learners' attention (Akçayır & Akçayır, 2017). Merging virtual elements with the likely surroundings plausibly generated situational interest and intense focus triggered by perceptual novelty (Renninger & Hidi, 2011), which the interviews suggest matured into more sustained engagement over the six weeks.

The effect on Relevance ($d = 1.24$) accords with experiential learning theory, in which concreteness of experience increases perceived personal relevance (Kolb, 2014). By handling virtual referents in everyday settings, learners appear to have perceived the direct usefulness of the words they were acquiring.

Confidence had the smallest effect ($d = 0.89$), and this pattern is informative in itself. Although AR offers scaffolded, low-anxiety practice, some learners struggled to operate the devices, experienced unstable marker tracking, or faced

cognitive overload (Bressler & Bodzin, 2013), which may have partly offset gains in self-efficacy. The qualitative accounts of an initial learning curve support this reading and point to simplified interface design and short orientation sessions as design priorities.

Satisfaction produced the largest effect of all ($d = 1.98$; $M = 4.48$ against 3.29). Within the ARCS framework, satisfaction derives from both intrinsic and extrinsic rewards (Keller, 2009). Hence, immediate feedback from responsive virtual objects, the game-like features embedded in the application, and the experience of discovery plausibly combined to generate positive affect. Notably, the interviews indicate that satisfaction was tied less to the technology itself than to visible evidence of progress.

5.3 Learner Experience

The qualitative themes both corroborate and extend the quantitative pattern. Learners consistently described the shift from passive reception to exploration that Kolb's cycle predicts and reported reduced anxiety and greater willingness to participate, a pattern consistent with a reduced affective filter and earlier findings on AR in language learning (Oto-Millera et al., 2025). Participants attributed improved recall to visual-spatial and embodied experience rather than memorisation, echoing existing evidence on AR-assisted vocabulary learning (Ma et al., 2025; M. Zhang et al., 2025).

Participants themselves distinguished initial novelty from sustained motivation, locating the latter in perceived improvement and growing self-efficacy. They also valued the autonomy that AR afforded, particularly the opportunity to rehearse without public exposure. Enjoyment was reported widely, and English lessons were frequently characterised as fun, consistent with earlier work on AR game-based language learning in elementary classrooms (Liao et al., 2023). Participants experienced technical difficulties and an initial learning curve, which they viewed as temporary and surmountable, supporting the view that technical readiness and infrastructure are common but manageable barriers to AR use in language teaching (Salehi, 2025).

Moreover, learners' reports of better vocabulary recall and contextual understanding were consistent with the adjusted between-group vocabulary advantage observed in the quantitative analysis. The convergence of self-reported learning experiences and post-test scores provides complementary evidence of educational promise, whereas interviews with 10 AR-group learners cannot establish which elements of the intervention led to the observed achievement difference. Further process-level studies could investigate whether repeated object manipulation, visual encoding, reflection, or contextual practice contributed most strongly to the achievement gains.

5.4 Theoretical and Practical Contribution

Taken together, the findings indicated that a pedagogically structured AR intervention is associated with gains in both measured vocabulary achievement and learner motivation. The significant adjusted achievement difference and the large ARCS effects suggest complementary cognitive and motivational benefits, although these outcomes should still be evaluated as distinct constructs. For a primary EFL population where disengagement is a recognised concern (Azar & Tanggaraju, 2020), these results provide an empirical rationale for integrating concrete AR experiences, reflection, conceptualisation, and active language use into vocabulary lessons. The present study contributes mixed-methods evidence from a multicultural primary setting that the literature has underrepresented (Akçayır & Akçayır, 2017; Qaddoura et al., 2026), while indicating the need to test durability, implementation fidelity, and generalisability across schools.

5.5 Implications for Practice

These findings suggest that AR should be integrated into EFL instruction as a complement to conventional teaching rather than a substitute for it. AR activities are most likely to be effective when designed around experiential learning principles, giving learners opportunities to interact with vocabulary, reflect on that interaction, connect visual representations with linguistic meaning, and apply new words in contextualised tasks. Given that confidence gains were the smallest of the four ARCS dimensions, teachers should provide brief orientation and guided practice before AR activities, especially for first-time users, and build in peer interaction, repeated attempts, scaffolding, and feedback.

At the policy and curriculum level, AR integration requires clear instructional objectives, aligned assessment, teacher professional development, technical support, and adequate digital infrastructure. Curriculum designers may also explicitly align AR activities with experiential learning principles and motivational frameworks, such as the ARCS model, while schools should establish routine equipment checks and alternative activities in case technology fails. Sustainable implementation depends on coordinated attention to pedagogy, teacher readiness, infrastructure, and learner experience, not on technology adoption alone.

5.6 Limitations and Future Research

Several limitations qualify these findings. The quasi-experimental design relied on intact classes without random

assignment, and the lack of a statistically significant pre-test difference does not establish equivalence of unmeasured student or classroom characteristics. Therefore, the study used an immediate post-test and did not examine longer-term vocabulary retention or motivational gains. The qualitative phase involved 10 purposively selected learners from the experimental group within a single school context, which limited the transferability of their accounts. The intervention depended on tablets, reliable marker recognition, and functioning classroom equipment. Findings may not generalise to other augmented-reality configurations. Vocabulary assessment primarily addressed receptive knowledge rather than productive language use, while motivation relied heavily on self-report. Although both groups followed equivalent instructional schedules and vocabulary content, differences in opportunities for individual practice, feedback quality, and implementation fidelity could not be entirely excluded. Consequently, the observed differences should not be attributed exclusively to AR technology.

As the same English teacher delivered both instructional conditions at each participating school, potential teacher expectancy effects and cross-condition contamination could not be entirely ruled out. Furthermore, students were nested within six intact classes across three participating schools, and differences in classroom composition, peer interactions, and school-level resources may have influenced the observed outcomes. The reported ANCOVA and MANCOVA analyses were conducted at the individual level without explicitly accounting for classroom clustering, potentially affecting the accuracy of standard errors and statistical significance. Hence, interpret the findings cautiously, as class- and school-level influences could not be fully separated from instructional effects. Future research should include more schools and intact classes, systematic monitoring of instructional fidelity, and appropriate cluster-aware analyses.

Future research should use longitudinal designs with delayed post-tests to assess whether vocabulary and motivational gains are sustained. Larger samples with more intact classes and schools, recorded teacher assignments, and cluster-aware analysis should be used to strengthen causal interpretation and generalisability. Other than that, comparative studies could contrast marker-based with markerless AR, individual with collaborative activity, and AR with virtual reality, gamification, and digital flashcards. Extending AR to productive skills, including speaking, pronunciation, writing, and communicative vocabulary use, would be valuable, as would examining moderators such as English proficiency, prior technological experience, learner characteristics, and the optimal duration and intensity of AR-based instruction, while complementing self-report with behavioural indicators of engagement.

6. Conclusion

This study determined that AR-based experiential learning was associated with significantly higher adjusted vocabulary achievement among grade six EFL learners in Arab schools in Malaysia than conventional instruction, $F(1, 117) = 75.68$, $p < .001$, partial $\eta^2 = .393$. The intervention was also associated with strong positive motivation outcomes, particularly Attention and Satisfaction, while interviewed learners described it as engaging, interactive, and enjoyable and reported improved vocabulary recall.

Thus, AR-supported experiential learning offers a promising instructional enhancement with demonstrated short-term advantages in reported achievement and motivation, rather than merely an engaging alternative to conventional teaching. When designed around experiential learning principles and supported by appropriate motivational and technical scaffolding, it can create more interactive and learner-centred language classrooms. By providing mixed-methods evidence from an under-researched multicultural primary population, this study extends current understanding of AR-supported EFL learning. Further work is needed to establish long-term retention and the conditions under which AR yields sustained benefits across diverse language-learning settings.

Acknowledgments

The authors would like to acknowledge the financial support from Universiti Teknologi Malaysia under UTM Encouragement Research Grant Scheme (UTMER) (Q.J130000.3853.42J25).

Author contributions

Mohammed Z. M. Qaddoura was responsible for study conceptualisation, research design, intervention implementation, data collection, formal analysis, and preparation of the original manuscript. Mohd Fadzil Abdul Hanid contributed to the study design, methodological supervision, interpretation of findings, and critical revision of the manuscript. Mohd Nihra Haruzuan Mohamad Said contributed to supervision and critical review of the manuscript. Huan Yik Lee contributed to the manuscript review and editing. All authors reviewed and approved the final version of the manuscript and agreed to be accountable for their respective contributions to the study.

Funding

The authors would like to acknowledge the financial support from Universiti Teknologi Malaysia under UTM Encouragement Research Grant Scheme (UTMER) (Q.J130000.3853.42J25).

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1–11. <https://doi.org/10.1016/j.edurev.2016.11.002>
- AlGerafi, M. A. M., Zhou, Y., Oubibi, M., & Wijaya, T. T. (2023). Unlocking the Potential: A Comprehensive Evaluation of Augmented Reality and Virtual Reality in Education. *Electronics*, 12(18), 3953. <https://doi.org/10.3390/electronics12183953>
- Altman, D. G., & Bland, J. M. (1995). Statistics notes: Absence of evidence is not evidence of absence. *BMJ*, 311(7003), 485. <https://doi.org/10.1136/bmj.311.7003.485>
- Aronson, B., & Laughter, J. (2015). The Theory and Practice of Culturally Relevant Education. *Review of Educational Research*, 86(1), 163–206. <https://doi.org/10.3102/0034654315582066>
- Azar, A. S., & Tanggaraju, D. (2020). Motivation in second language acquisition among learners in Malaysia. *Studies in English Language and Education*, 7(2), 323–333. <https://doi.org/10.24815/siele.v7i2.16506>
- Bamanger, E. (2025). Augmented Reality in EFL Education: Transforming High School Student Achievement Through Interactive Learning Tools. *Journal of Technology Studies*, 50(1), 30–41. <https://doi.org/10.21061/jts.439>
- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Preservice Teachers' Knowledge and Attitudes toward Digital-Game-Based Language Learning. *Education Sciences*, 12(3), 182. <https://doi.org/10.3390/educsci12030182>
- Binhomran, K., & Altalhab, S. (2021). The impact of implementing augmented reality to enhance the vocabulary of young EFL learners. *The JALT CALL Journal*, 17(1), 23–44. <https://doi.org/10.29140/JALTCALL.V17N1.304>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Bressler, D. M., & Bodzin, A. M. (2013). A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. 505–517. <https://doi.org/10.1111/jcal.12008>
- Chan, C. K. Y. (2022). Assessment for Experiential Learning. In *Directory of Open access Books (OAPEN Foundation)*. OAPEN Foundation. <https://doi.org/10.4324/9781003018391>
- Chang, H. Y., Binali, T., Liang, J. C., Chiou, G. L., Cheng, K. H., Lee, S. W. Y., & Tsai, C. C. (2022). Ten years of augmented reality in education: A meta-analysis of (quasi-) experimental studies to investigate the impact.

- Computers & Education*, 191, 104641. <https://doi.org/10.1016/J.COMPEDU.2022.104641>
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Creswell, J. W. (2018). *Designing and Conducting Mixed Methods Research*, 3rd ed. https://catalog.maranatha.edu/index.php?p=show_detail&id=51811
- Daud, W. A. A. W., Teck, W. K., Ghani, M. T. A., & Ramli, S. (2020). M-Learning Boost Students' Motivation in Learning Arabic Language Proficiency for Elementary Level. *Universal Journal of Educational Research*, 8(10), 4384–4392. <https://doi.org/10.13189/ujer.2020.081004>
- Doughty, C. J., & Long, M. H. (2003). *The Scope of Inquiry and Goals of SLA* (pp. 1–16). <https://doi.org/10.1002/9780470756492.ch1>
- Doyle, N., McDowall, A., Randall, R., & Knight, K. (2022). Does it work? Using a Meta-Impact score to examine global effects in quasi-experimental intervention studies. *PLoS ONE*, 17(3). <https://doi.org/10.1371/journal.pone.0265312>
- Elena V. Frolova, Olga V. Rogach, & Tatyana M. Ryabova. (2020). Digitalization of Education in Modern Scientific Discourse: New Trends and Risks Analysis. *European Journal of Contemporary Education*, 9(2). <https://doi.org/10.13187/ejced.2020.2.313>
- Ghani, M. T. A., Shalikin, N. H. M., Ramli, S., Zakaria, Z. M., Daud, W. A. A. W., Amiruddin, A., & Ahmad, M. A. (2025). Enhancing Arabic communication proficiency in Malaysian among Malaysian students: Barriers, pedagogical strategies, and environmental influences. *Edelweiss Applied Science and Technology*, 9(7), 2078–2089. <https://doi.org/10.55214/2576-8484.v9i7.9093>
- Godwin-Jones, R. (2004). Language in Action: From Webquests to Virtual Realities. *Language Learning & Technology*, 8(3), 9–14. <https://doi.org/10.64152/10125/25246>
- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275–285. <https://doi.org/10.1016/j.susoc.2022.05.004>
- Hung, Y.-H., Chen, C., & Huang, S. (2016). Applying augmented reality to enhance learning: a study of different teaching materials. *Journal of Computer Assisted Learning*, 33(3), 252–266. <https://doi.org/10.1111/jcal.12173>
- Jaliyya, F. G., & Idrus, F. (2017). EFL Students' Attitudes and Perception towards English Language Learning and Their English Language Proficiency: A Study from Assa'adah Islamic Boarding School, Indonesia. *Journal of Education and Learning (EduLearn)*, 11(3), 219–228. <https://doi.org/10.11591/edulearn.v11i3.4621>
- Jia, W., Pack, A., Guan, Y., Zhang, L., & Zou, B. (2025). The influence of game-based learning media on academic English vocabulary learning in the EFL context. *Computer Assisted Language Learning*, 38(5–6), 1341–1365. <https://doi.org/10.1080/09588221.2023.2276800>
- Keller, J. M. (2009). *Motivational Design for Learning and Performance*. <https://doi.org/10.1007/978-1-4419-1250-3>
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development* (2nd ed.). Pearson FT Press.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Larsen-Freeman, D., & Freeman, D. (2008). Language Moves: The Place of “Foreign” Languages in Classroom Teaching and Learning. *Review of Research in Education*, 32(1), 147–186. <https://doi.org/10.3102/0091732x07309426>
- Liao, C.-H. D., Wu, W. V., Gunawan, V., & Chang, T.-C. (2023). Using an Augmented-Reality Game-Based Application to Enhance Language Learning and Motivation of Elementary School EFL Students: A Comparative Study in Rural and Urban Areas. *The Asia-Pacific Education Researcher*, 33(2), 307–319. <https://doi.org/10.1007/s40299-023-00729-x>
- Ma, C., Haire, G., Smith, S., & Newbutt, N. (2025). Enhancing English as foreign language education with virtual reality and augmented reality in K–12 classrooms: A systematic review from 2019 to 2023. *Technology in Language Teaching & Learning*, 7(3), 102535. <https://doi.org/10.29140/tl.v7n3.102535>
- Mohanasundaram, S. S. T., BV, P., & Harsha, H. (2024). Selecting the right sample size: Methods and considerations for social science researchers. *International Journal of Business and Management Invention*, 13(7), 152–167.
- Nattida, P. (2021). Culture teaching practices of Lower Secondary School ELT teachers from the Global English perspective: A qualitative case study in Thailand. In *ePrints Soton (University of Southampton)*. <https://doi.org/10.5258/soton/t0032>
- Oto-Millera, N., Pellicer-Ortín, S., & Bustamante, J. C. (2025). Augmented Reality in English Language Acquisition Among Gifted Learners: A Systematic Scoping Review (2020–2025). *Applied Sciences*, 15(21), 11487.

<https://doi.org/10.3390/app152111487>

- Park, M., & Sung, Y.-K. (2013). Teachers' perceptions of the recent curriculum reforms and their implementation: what can we learn from the case of Korean elementary teachers? *Asia Pacific Journal of Education*, 33(1), 15–33. <https://doi.org/10.1080/02188791.2012.756391>
- Parmaxi, A., & Demetriou, A. A. (2020). Augmented reality in language learning: A state-of-the-art review of 2014–2019. *Journal of Computer Assisted Learning*, 36(6), 861–875. <https://doi.org/10.1111/JCAL.12486>
- Qaddoura, M. Z. M., Hanid, M. F. A., Said, M. N. H. M., & Li, Y. S. (2026). The Impact of Augmented Reality on Motivation, Engagement and Language Proficiency in EFL Education: A Systematic Review. *International Journal of Learning, Teaching and Educational Research*, 25(2), 1–24. <https://doi.org/10.26803/ijlter.25.2.1>
- Rass, R. A. (2024). Palestinian Arab Students Writing Research Papers: Challenges and Remedies. *Magiâ INNO.*, 6(2). <https://doi.org/10.24833/2949-6357.2024.rar.2>
- Renninger, K. A., & Hidi, S. (2011). Revisiting the Conceptualisation, Measurement, and Generation of Interest. *Educational Psychologist*, 46(3), 168–184. <https://doi.org/10.1080/00461520.2011.587723>
- Salehi, N. (2025). Teaching vocabulary with augmented reality: voices from junior high school English teachers. *Discover Education*, 4(1). <https://doi.org/10.1007/s44217-025-00673-6>
- Saud, A. M., Ghalib, Mohd. F. Md., & Bakar, R. A. (2023). A STRUCTURED REVIEW OF MOBILE AUGMENTED REALITY FOR LANGUAGE INSTRUCTION AND LEARNING. *International Journal of Modern Education*, 5(19), 78–98. <https://doi.org/10.35631/ijmoe.519006>
- Wen, Y. (2020). Augmented reality enhanced cognitive engagement: designing classroom-based collaborative learning activities for young language learners. *Educational Technology Research and Development*, 69(2), 843–860. <https://doi.org/10.1007/s11423-020-09893-z>
- Yuen, S. C.-Y., Yaoyuneyong, G., & Johnson, E. (2011). Augmented Reality: An Overview and Five Directions for AR in Education. *Journal of Educational Technology Development and Exchange*, 4(1). <https://doi.org/10.18785/jetde.0401.10>
- Zhang, M., Hashim, H., & Yunus, M. M. (2025). Analysing and comparing augmented reality and virtual reality assisted vocabulary learning: a systematic review. *Frontiers in Virtual Reality*, 6. <https://doi.org/10.3389/frvir.2025.1522380>
- Zhang, Y., & Miao, Z. (2025). Enhancing EFL Learners' Engagement and Motivation Through Immersive Technologies: The Role of Artificial Intelligence, Augmented Reality, Virtual Reality, and Mobile Applications. *European Journal of Education*, 60(2). <https://doi.org/10.1111/ejed.70128>