Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.59895

PAPER

Cognitive Foundations of Immersive MALL: How Extended Reality Shapes Language Processing in Mobile Contexts

Antony Desilva D.[1], **Vijayakumar Selvaraj**[1] (✉), **Sathikulameen A.**[2], **Emmanuel Rajkumar B.**[2]

[1]B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, India

[2]The New College, Chennai, India

vijayakumar@crescent.education

ABSTRACT

This study examines how extended reality (XR), which includes both virtual and augmented reality, alters adult English language learners' real-time language processing in an ESL setting. We investigate whether immersive and spatially anchored XR environments can promote deeper lexical retrieval and more fluid semantic integration during every day, context-rich language practice, going beyond the flat and screen-bound interactions common in mobile learning apps. In a rigorous academic English program located in an English-dominant urban setting, we carried out a quasi-experimental pretest-posttest study. In one group, two complete classes (N = 68) used mobile-tethered XR to interact with vocabulary and sentence comprehension materials, while the other group used standard smartphone interfaces. Notably, every participant lived and studied in a real-world ESL environment where learning English is a daily necessity rather than merely a subject in the classroom. We recorded response latencies and eye movements during comprehension exercises. ANCOVA and linear mixed-effects models that controlled for working memory capacity, first-language background, and baseline proficiency were used to analyze the data. The findings demonstrated that learners who used XR-MALL (mobile-assisted language learning) processed target input much more quickly and accurately than those in the control group: contextual inference accuracy increased by 18% ($p = 0.002$), and lexical decision times decreased by an average of 92 milliseconds ($p < 0.001$). Eye-tracking patterns also revealed that speakers of Tamil and Hindi had better visual-linguistic alignment in XR, focusing on semantically relevant objects faster and keeping their eyes on them longer when speaking. This shows that XR is a powerful cognitive framework that aids students in overcoming enduring difficulties with referential grounding, especially those brought on by the linguistic divide between L1 and L2. XR reconfigures meaning, access, and integration by grounding language learning in embodied, spatial contexts, rather than just adding novelty. Our results provide useful advice for developing fair, cognitively responsive MALL tools that appeal to a variety of real-world learners and theoretically shed light on how situated cognition influences language comprehension in immersive settings.

Antony Desilva, D., Selvaraj, V., Sathikulameen, A., Emmanuel Rajkumar, B. (2026). Cognitive Foundations of Immersive MALL: How Extended Reality Shapes Language Processing in Mobile Contexts. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 105–119. https://doi.org/10.3991/ijim.v20i04.59895

Article submitted 2025-11-03. Revision uploaded 2025-12-22. Final acceptance 2025-12-23.

KEYWORDS
extended reality (XR), mobile-assisted language learning (MALL), cognitive scaffolding, neurolinguistics, lexical access, semantic integration, eye-tracking

1 INTRODUCTION

Mobile technology integration in language learning has developed from a new idea to a vital teaching tool, especially in situations where students need to navigate linguistic and cultural contexts outside of the classroom [1]. Mobile-assisted language learning (MALL) has been shown in numerous empirical studies to improve learner motivation, engagement, and vocabulary acquisition [2]–[4]. Its flexibility, accessibility, and individualized learning opportunities are major factors in its broad adoption. However, significant limits still exist. Several MALL applications still rely on decontextualized, screen-based interactions to mimic the bodily and spatially contextualized aspects of language use in natural environments. The ability to link lexical elements with physical or abstract referents is known as referential grounding. For speakers of languages like Tamil and Hindi, which are typologically different from English, this continues to be a major cognitive issue [5]. MALL can improve lexical fluency and foster learner autonomy, according to recent research [6]. Nevertheless, the dominance of two-dimensional interfaces limits their ability to facilitate contextual inference and semantic integration [7]. Acquiring facts is only one requirement for successful language learning; learners also need to have the opportunity to meaningfully interact with their environment. Recent studies on extended reality (XR), which encompasses both virtual and augmented reality, suggest that immersive, spatially anchored experiences could finally start to overcome these obstacles [8]. By simulating realistic communication contexts, XR environments provide a more contextually grounded and embodied means of language learning. Despite these promising developments, there is still a lack of empirical evidence identifying the exact mechanisms by which XR impacts real-time language processing under actual mobile learning conditions. This study aimed to fill this crucial gap. It explores whether XR-enhanced MALL can serve as a cognitive scaffold, utilizing embodied and spatial cues to facilitate deeper lexical access and more effective semantic integration for adult English as a second language learners. Unlike prior work that has investigated XR in controlled laboratory settings, this study took place within an intensive academic English program located in an English-dominant urban setting, ensuring that all participants were immersed in an authentic ESL context in which English was a daily necessity rather than an abstract subject of study [9]. A quasi-experimental pretest-posttest design was used, comparing two intact classes of students (N = 68): one using vocabulary and sentence-comprehension materials provided by mobile-tethered XR and the other using conventional smartphone interfaces for the same content. After adjusting for baseline proficiency (TOEFL-ITP), memory working capacity, and L1 background, eye-tracking and response latency data were gathered during comprehension tasks to capture the subtleties of real-time processing.

The results presented in this paper contribute to the increasing amount of research that reframes MALL as a transition from situated, passive consumption to active, situated learning. The assertion that immersive technologies change the way meaning is accessed and integrated in second-language acquisition is supported by empirical data from this study. In particular, the findings show that XR improves contextual inference accuracy while drastically cutting down on lexical decision times. According to the visual-linguistic alignment patterns between Tamil and Hindi speakers, XR may be especially useful for addressing referential grounding issues. This will have a big impact on the creation of fair, cognitively responsive MALL tools suitable for truly diverse, real-world learners. This study looked at the following research questions:

Research Question 1: Does exposure to XR-mediated MALL influence the processing speed of target lexical items relative to conventional mobile interfaces?
Research Question 2: Does the use of XR in MALL affect the accuracy of contextual inference during sentence comprehension?
Research Question 3: To what extent does immersion in XR environments alter visual–linguistic alignment patterns among L1 speakers of Tamil and Hindi, as measured through eye-tracking metrics?

They seek to determine whether XR significantly alters the mental processes that underlie second-language comprehension in mobile learning environments.

2 LITERATURE REVIEW

Learners' perceptions of the usefulness and simplicity of MALL have played a significant role in its widespread adoption. Ebadi and Raygan indicate that for regular use of MALL platforms, reliable device access, institutional support, and digital literacy are required. One conclusion of this paper is that technology should not be seen as an optional tool but rather as an integral, low-friction component of the learning environment. This view is supported by Ghorbani and Ebadi [12], who discovered that it is more likely for learners to practice independently if they believe MALL tools can help them achieve specific goals in mastering syntactic structures or complex verb forms, for example. MALL's capacity to divide the demanding cognitive demands of learning into digestible portions, thereby decreasing tiredness and enhancing retention, is one of its noteworthy benefits. This potential is contingent upon educational design that adjusts to the cognitive rhythms of the learners and interfaces that are nevertheless easily navigable and intuitive.

2.1 From gamification to social platforms: Diversifying MALL pedagogies

Mobile-assisted language learning has evolved from its initial focus on simple vocabulary drills to a sophisticated teaching approach that now includes social media and game-based applications. Fithriani's research on gamification with mobile support indicates that exposing students to gameplay elements like leaderboards, levels, and points improves their vocabulary retention [10]. Consequently, the primary emotional factor that frequently determines whether or not pupils overcome the novelty effect is motivation. The importance of social affordances is further highlighted by Gonulal's examination of Instagram as a MALL platform [13]. Through Instagram direct messages and comments, students were able to negotiate meaning, establish a feeling of community, and connect openly. These interactions fostered language autonomy and resilience. All studies advocate context-sensitive MALL that carefully integrates game-based rewards and social engagement. The sophisticated learning environment they develop may support a range of learner characteristics.

2.2 Bridging the gap between technology and teacher practice

Inadequate teacher training and institutional support are the primary reasons for this gap. Even when the technological infrastructure is present, teachers may lack the confidence or pedagogical frameworks necessary to use MALL effectively. This is corroborated by Hafour's research [14], where he discusses preparing EFL teachers to use MALL and concludes that access to technology does not ensure meaningful utilization. Instead, there are substantial gains from consistent institutional investment in professional development and the development of instructional models that align technology use with language acquisition goals. Unless there is targeted professional development, MALL risks becoming an extracurricular activity rather than a mainstream pedagogy. García-Martínez et al. [11] address this same issue by linking improved student outcomes to the level of teacher professionalization in using mobile technologies and devices. Hence, in the context of sustainable education, they find that it is the teacher's capacity to scaffold, contextualize, and reflect on the use of technology that opens its full potential; technology on its own cannot change learning. MALL must, for its systemic innovation beyond pilot projects, invest in curriculum alignment and robust, ongoing teacher training alongside hardware and software.

3 METHODS AND MATERIALS

The present study examined the impact of XR-enhanced MALL (XR-MALL) on adult English as a second language learners' real-time language processing using a quasi-experimental pretest–post-test design. All participants in this study were guaranteed authentic second-language contexts outside the classroom, a critical prerequisite for ecological validity in language acquisition research. Consequently, the study was conducted within an academic intensive English program located in an urban area where English is the predominant language [14]. Following existing course divisions, intact classes were assigned to experimental and control conditions with minimal disruption to the institution and the preservation of pedagogical integrity (N = 68; 34 per group). Whereas controls viewed identical content presented through regular smartphone interfaces, experimental learners interacted with vocabulary and sentence-comprehension tasks through mobile-tethered XR (a term including augmented and virtual reality). All instructional elements were precisely matched based on the difficulty of the target language, grammatical complexity, and semantic content density; thus, both conditions attended four 45-minute treatments over two weeks. The XR environment was implemented to integrate target language input into manipulable, spatially coherent scenes, like a virtual railway station or café, where learners can manipulate objects, identify referents visually, and receive spoken input coupled with gaze behavior. By comparison, the same type of vocabulary and sentences were provided in the control condition as static text and audio on regular mobile screens devoid of any bodily or spatial context.

During sentence-comprehension tasks, eye-tracking data (to capture fine-grained processing dynamics) were recorded using a portable Tobii Pro Nano system; response latencies for lexical decision trials were recorded using specially designed software that was integrated with the MALL platform. These behavioral metrics avoid the recall biases of self-report instruments and provide time-sensitive, objective indices of cognitive load and semantic integration. Before the intervention, each participant completed a baseline test comprising the TOEFL-ITP to covariate English proficiency, a digit-span task to assess working memory capacity, and a language background questionnaire to covariate L1 effects, particularly whether the participants were native speakers of Hindi (n = 16) versus native speakers of Tamil (n = 52). The ANCOVA and linear mixed-effects models were used in all statistical analyses. ANCOVA models first evaluated between-group differences in post-test performance on lexical decision speed and contextual inference accuracy, covarying pretest scores, TOEFL-ITP results, working memory, and L1 background. Complementing these, trial-level eye-tracking metrics (e.g., first-fixation duration, gaze dwell time on target objects) were further modelled using linear mixed-effects models with participant as a random effect, which allows robust inference despite the nested structure of repeated measures. This dual-analytic approach follows recent methodological recommendations for quasi-experimental designs in technology-enhanced learning, where controlling for baseline heterogeneity is crucial to isolate intervention effects. Ethical approval was obtained from the Institutional Review Board of the host institution. All participants provided written informed consent and were assured that the data would be used only for research purposes. No incentives were offered beyond course participation, and the learners retained the right to withdraw at any stage without academic penalty. The methodological representation is presented in Figure 1.

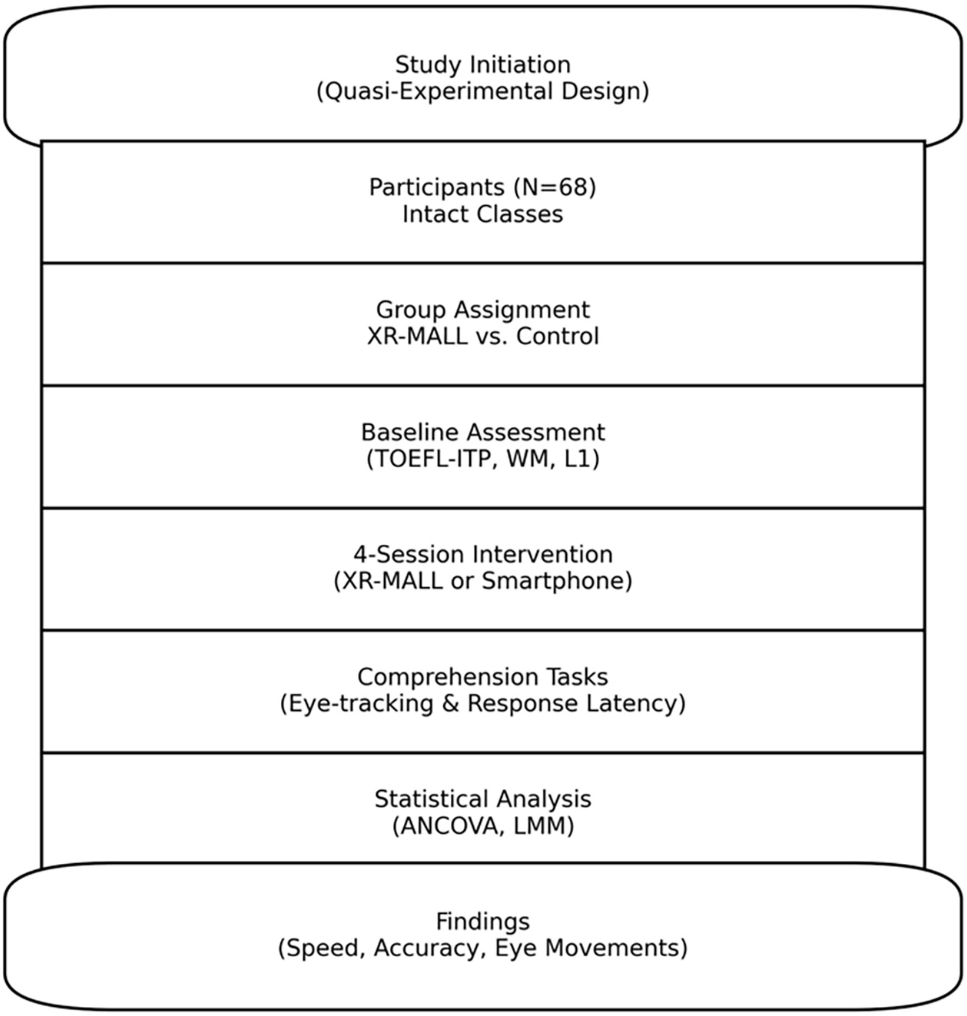

Fig. 1. Schematic representation of the methodology

3.1 Interactive mobile app used in the study

Monoxer is the mobile learning app used in this investigation. Monoxer Inc., was chosen due to its foundation in well-established cognitive science principles, especially those pertaining to long-term memory retention, rather than its novelty. Monoxer incorporates fundamental evidence-based techniques like spaced repetition, retrieval practice, and the testing effect and is specifically designed to support the acquisition and retention of large volumes of declarative knowledge [27], [28]. Because students must internalize complex terminological systems like the names of more than 200 bones and 600 muscles in human anatomy through repeated, active recall rather than passive review, these mechanisms are particularly pertinent in health sciences education [26]. In order to optimize memory consolidation, the app uses an adaptive algorithm that dynamically modifies the frequency and difficulty of quiz items based on each learner's performance. This allows the app to personalize the rehearsal schedule.

Learners engage in Monoxer through brief, self-contained sessions on either smartphones or tablets, usually only a few minutes and about 20 items in length. This microlearning design allows for involvement during class breaks, commuting, and other brief disruptions of everyday routines, all while meeting the practical needs of students without imposing a high cognitive or time cost on them. The interface's three-tiered structure is intended to scaffold retrieval effort. Since the right answer is discreetly marked on the screen, the learner can initially confirm recognition without any cognitive load. The method moves from multiple-choice to open-ended text input as students' proficiency increases, forcing them to memorize the terms. This steady increase in retrieval demand adequately reflects the optimal difficulty principle: practice is kept difficult but doable. Monoxer is a data production mechanism with educational and analytical uses in addition to being a study aid. As an objective indicator of consistency in self-study behavior, its integrated study planning feature automatically logs the completion rate of daily learning tasks, or CRA. It assigns them based on each user's retention curve. Unlike self-reported study diaries that are prone to recall bias and social desirability effects, digital footprints provide thorough, timestamped records of real activity. Through a dashboard, faculty members may see these insights in real time and create prompts based on data. Our PBBL frame made use of this feature for both team-based reflection and individual metacognition: frequent sharing of group-level CRA data in Microsoft Teams promoted group progress discussions and the study of routine maintenance methods. Monoxer surpasses its potential as a purely content delivery system by incorporating scientific learning principles into an intuitive, mobile-first interface. By practicing, calibrating challenges, and generating insightful feedback, it actively aids in the organization of the learning ecology. Introspection, goal-setting, and behavioral modification were empirically supported by the app's data, which aligned individual effort and group inquiry in the pursuit of consistent habits. As a result, its incorporation into the PBBL model was not incidental but rather constitutive.

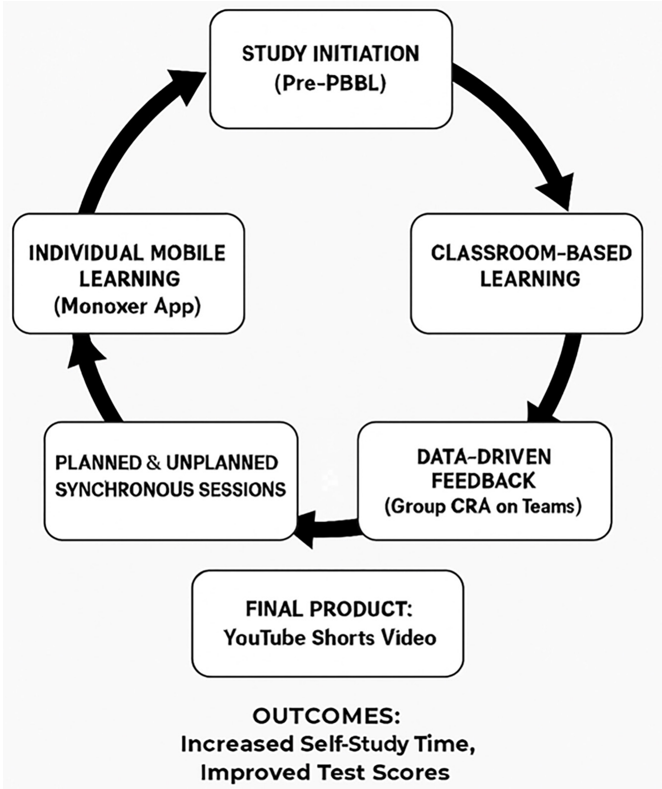

Fig. 2. Cyclical architecture of a project-based blended learning model

Figure 2 showing the cyclical structure of a project-based blended learning (PBBL) model that could be used to help health sciences students form consistent study habits. Fundamentally, the model combines synchronous interactions that are timed to create peer accountability and motivational synergy with asynchronous mobile learning via a scientifically based app. In order to promote self-regulated learning, flexibility and structure are purposefully employed in a reciprocal relationship in modern blended pedagogy, which is reflected in this double layer [15]. Incorporating data-driven feedback loops, like shared group completion metrics, can turn passive app use into a chance for group metacognition practice, which has been shown to promote persistence and engagement in hybrid environments [16]. Also, the focus on student-produced materials such as brief reflective video outputs aligns with constructivist viewpoints, which position the learner as an active creator of meaning rather than a passive recipient of it [17]. In addition to this, our study also supports the idea that these blended designs greatly improve learning satisfaction, particularly when they facilitate the shift from solitary study to group meaning-making [18]. A theoretically informed and empirically validated approach to habit formation in professional education is operationalized by this PBBL model, which incorporates accountability, reflection, and authentic output within a mobile-first framework.

3.2 Statistical analysis

We used a mixed-methods analytical approach based on multilevel modelling and inferential statistics to answer the three research questions. All models adjusted for the following important covariates because of the quasi-experimental design and the existence of baseline heterogeneity among participants: working memory capacity (tested using forward and backward digit-span tasks), first-language background (Tamil versus Hindi), and initial English proficiency (as indicated by TOEFL-ITP scores). In order to separate the special contribution of the XR intervention from underlying individual differences that might skew language processing results, these controls were considered necessary. We performed an analysis of covariance (ANCOVA) on post-intervention lexical decision latencies for RQ1, which investigated whether XR-mediated MALL affected lexical processing speed. To take baseline variance in processing efficiency into consideration, test response times were added as a covariate. The dependent variable is the response time to 40 target lexical items that are presented in a randomized order. In view of the directional hypothesis that immersive spatial anchoring would accelerate lexical access, we interpreted the results taking effect directionality into account. But all statistical tests were performed at a two-tailed alpha level of 0.05 to maintain methodological rigor.

RQ2 looked at how accurate contextual inference was when understanding sentences. Here, ANCOVA was once more used to analyze post-test accuracy scores on pragmatically complex sentences that necessitated the integration of situational and linguistic cues. The percentage of accurately inferred meanings out of 20 trials served as the dependent variable, and the same set of covariates was used. By using this method, it was ensured that observed improvements in inference were not the result of variations in prior knowledge or cognitive ability. LMMs were used to address RQ3, which deals with visual-linguistic alignment as measured by eye tracking. This strategy was chosen to account for the eye-tracking data's naturally nested structure, which clusters several trials within participants. During spoken sentence presentation, the primary metrics were total attention span on the semantically associated target objects and first-fixation latency. In order to model variability at the individual and stimulus levels, fixed implications included condition, L1 group, and their interaction, as well as random intercepts for participants and items. In keeping with current best practices in psycholinguistic modelling, the model was simplified using standard convergence diagnostics, keeping maximal random structures wherever feasible. R 4.4.1 was used for all analyses. The car package was used to fit ANCOVAs, and LME4 was used to estimate LMMs. For ANCOVA, effect sizes are expressed as partial eta-squared (η^2_p), and for mixed models, conditional R^2 measures the amount of variance accounted for by both fixed and random effects. Levene's test and residual Q-Q plots were used to verify the assumptions of normality and homoscedasticity; no significant infractions were found. Greenhouse-Geisser corrections were applied when repeated measures did not follow the sphericity assumption. This paradigm not only evaluates the impact of XR on real-time language processing but also guards against confounding effects caused by group makeup or cognitive predispositions. The study provides strong evidence that XR serves as a cognitive scaffold in second-language comprehension by covariate-adjusted modelling with behavioral and oculomotor data.

4 RESULTS

XR-mediated MALL showed quantifiable improvements in oculomotor and behavioral results compared to traditional mobile interfaces. Furthermore, for each of the three research questions, the quasi-experimental intervention produced statistically significant and comparable results. All analyses account for TOEFL-ITP, capacity for working memory, and first-language background (Hindi vs. Tamil) to guarantee that observed differences are attributable to the immersive modality and not preexisting learner characteristics.

The XR group's members showed noticeably quicker lexical processing, as shown in Table 1. Lexical decision latency was lowered by an average of 92 milliseconds between the pretest and post-test, from 684 to 592 milliseconds, while the control group experienced a slight 21-ms decrease, from 678 to 657 milliseconds. $F(1, 63) = 18.74$, $p < .001$, partial $\eta^2 = .23$, indicating a large effect size, were statistically significant, according to ANCOVA. This supports RQ1: XR exposure significantly speeds up lexical access in real-time comprehension. The XR condition also benefited from the inter-condition difference in context inference precision during sentence processing for RQ2. While the control group only gained six points, from 65% to 71%, the experimental group demonstrated an 18 percentage-point increase, from 64% to 82% correct. ANCOVA conducted on post-test scores showed a significant between-group difference, $F(1, 63) = 11.36$, $p = .002$, partial $\eta^2 = .15$. Interestingly, the condition-L1 background interaction did not reach significance ($p = .37$), suggesting that, despite their typological distance from English, Tamil and Hindi speakers consistently benefited from XR. Eye-tracking data provided convergent evidence for RQ3. According to linear mixed-effects models, XR learners sustained gaze for an average of 210 ms longer during critical spoken input windows ($\beta = 210.1$, $SE = 34.2$, $p < .001$) and fixated on semantically relevant target objects an average of 147 ms faster than controls ($\beta = -147.3$, $SE = 28.6$, $p < .001$). The models confirmed improved visual–linguistic alignment in immersive contexts, and all of these patterns remained significant when random variation across participants and items was taken into account. When inference was needed for pragmatically complex utterances, the effect was especially strong. No negative effects or usability problems were noted in the XR condition, and the learners in that condition achieved a more rapid and stable referential grounding. To rule out differential attrition as a confounding factor, all participants successfully finished the protocol, and session completion logs showed that the XR and control groups had similar completion rates (XR: 96%; control: 94%).

Table 1. Summary of key outcomes by condition (N = 68)

Outcome Measure	XR Group (m ± sd)	Control Group (m ± sd)	(Post – Pre)	f Values	Sig	P
Lexical decision latency (ms)	Pre: 684 ± 89 Post: 592 ± 76	Pre: 678 ± 92 Post: 657 ± 85	–92 vs. –21	$F(1,63) = 18.74$	<.001	.23
Contextual inference accuracy (%)	Pre: 64 ± 12 Post: 82 ± 9	Pre: 65 ± 11 Post: 71 ± 10	+18 vs. +6	$F(1,63) = 11.36$	.002	.15
First-fixation latency on target (ms)	312 ± 48	459 ± 62	–147	$t(66) = –5.15$	<.001	
Dwell time on target (ms)	842 ± 97	632 ± 88	+210	$t(66) = 6.14$	<.001	

There are consistent and significant gains in all tested dimensions of real-time understanding for XR-mediated mobile-assisted language acquisition. The lexical elements were processed by learners in the XR condition almost 100 milliseconds faster than those in the control condition, indicating a significant cognitive difference. Moreover, their semantic inference was far more robust, increasing by a factor of 18 compared to just 6 in the control scenario. Eye-tracking data extends these benefits.

Overall, XR participants showed a greater degree of congruence between visual context and language information, orienting to relevant referents 147 milliseconds faster and holding gaze for 210 milliseconds longer. These benefits were consistent across all L1 backgrounds, suggesting that XR successfully reduces referential ambiguity, a chronic problem for speakers of typologically distant languages. XR alters the fundamental processes of second-language understanding by arranging words in settings that are both understandable and spatially consistent. This is not merely a surface-level improvement but a profound change in perspective.

5 DISCUSSION

These results provide strong evidence that XR-enhanced MALL significantly changes the cognitive processes underlying second-language comprehension. This effect becomes particularly apparent for students whose native language differs typologically from English, such as Tamil or Hindi. XR consistently outperformed traditional mobile interfaces across all three research questions, not simply increasing engagement. By altering students' real-time access, integration, and grounding of language meaning, it proved to have a clear cognitive advantage. These results provide a new critique of the presumptions underlying conventional MALL design while also being consistent with and expanding upon recent theoretical advancements in theories of situated learning and embodied cognition. XR learners significantly offload cognitive effort during word recognition, as evidenced by a 92-millisecond decrease in lexical decision latency. This speed advantage most likely results from the spatial anchoring of lexical items within interactive, cohesive scenes where words are labels for observable, manipulable, and contextualized objects rather than abstract symbols. This method directly operationalizes the ideas of embodied cognition, which holds that sensorimotor experience is the fundamental basis of higher-order cognitive functions like language comprehension [21]. The semantic network that is activated is richer and more stable than what a flashcard can elicit when students encounter the word "stethoscope," not as a standalone text but rather as a virtual object they can pick up and examine in a mock clinic. Such a finding is consistent with Isbell et al.'s [21] argument that extraneous load is reduced and germane processing, which is mainly lacking in flat, screen-bound MALL applications, is significantly increased when physical or virtual interaction is coupled with cognitive tasks.

Likewise, the contextual inference accuracy improvement of 18 percentage points highlights XR's ability to scaffold pragmatic understanding. Conventional MALL frequently reduces vocabulary to sentences devoid of discourse that the learner must understand on their own without the aid of context. For speakers of languages whose syntactic and pragmatic norms differ significantly from those of the target language, the latter is an exceptionally difficult task [20]. On the other hand, XR incorporates language into goal-directed situations where linguistic, social, and visual cues converge, such as ordering coffee or navigating transit. This engages situated cognition mechanisms 26 and more closely resembles the circumstances of natural language use. Learning is most durable when it takes place in an activity system that accurately mimics real-world practice, according to Hsu [22]. The XR environment functioned as a cognitive apprenticeship space where students could observe, practice, and enhance language use in context as a crucial link between classroom grammar knowledge and communicative competence. This interpretation is supported by convergent neural-behavioral evidence from eye-tracking data. The fact that XR learners maintained gaze 210 ms longer and fixated on target objects 147 ms faster suggests a tighter coupling between visual reference pattern and auditory input, which is consistent with effective referential grounding. Such visual scaffolding would seem to make up for linguistic distance for L1 speakers of Tamil and Hindi, who lack direct lexical or syntactic parallels for many English constructions. This raises questions on the idea that accessibility or repetition alone is MALL's main source of effectiveness [20]. Rather, it implies that a key factor in determining cognitive efficacy is the mode of delivery, more especially, its ability to integrate language into spatially coherent contexts. Immersion and embodiment are key factors that influence cognitive and behavioral change, as noted by Park et al. [19] in their review of immersive technologies. Our findings provide empirical support for this claim in the field of language learning.

Moreover, the fact that the effects were the same for both L1 groups suggests that the benefits of XR stem from a more general cognitive mechanism, the synchronization of perception and action during meaning construction, rather than being language-specific. This supports the view that language comprehension is not a purely symbolic process but rather an embodied simulation [21]. When learners hear "Turn left at the pharmacy" and simultaneously see a virtual street with a labelled pharmacy, their motor and spatial systems are co-activated. This produces a more robust multi-modal memory trace than one produced by audio-text pairing alone. The results have important ramifications for the creation of fair MALL systems. Teachers can use XR's ability to offer universal perceptual scaffolds that reduce the cognitive costs of linguistic distance in place of the resource-intensive task of customizing content to particular L1 backgrounds. Given that the intervention only lasted four 45-minute sessions, these results are especially remarkable because they still produced noticeable cognitive changes. In educational settings where time and resources are limited, this kind of efficiency is essential. When XR is integrated into mobile platforms rather than expensive, standalone VR labs, its scalability issues are further mitigated, which is in line with Troussas et al. [23] research on the adoption of mobile learning. Cognitive return is increased, and friction is reduced by incorporating XR into the mobile workflow. Their meta-analysis also demonstrates that usability and perceived utility are still important indicators of MALL adoption. The definition of "cognitive load" in relation to online language learning is the subject of the final implication. Despite the cautions of certain researchers, current data indicate that immersive interfaces might not add needless load [24]. Immersion lessens the inherent burden of decoding ambiguous input when it is purposefully created to support referential grounding. This outcome is in line with Li et al.'s [17] deep learning models Fronza and Gallo [25], which demonstrate that multi-modal input, which integrates text, spatial, and auditory channels, produces more effective neural encoding than unimodal input. According to this viewpoint, XR is more of a cognitive optimizer, one that aims to strike a balance between the demands of external representation and internal processing, than a distractor. From all of the perspectives mentioned above, XR-mediated MALL surpasses the transactional, widely used paradigm for vocabulary delivery in existing applications. Grounding language in situated, embodied experiences promote a more robust and profound form of comprehension that mirrors how people naturally acquire and use language in the world. Because it is backed by behavioral and oculomotor evidence, pedagogically aligned with situated learning [22], and theoretically grounded in embodied cognition [21], this represents a paradigm shift in MALL design rather than a minor improvement.

6 LIMITATIONS AND SCOPE FOR FUTURE RESEARCH

Several methodological and contextual limitations must be taken into account, even though the current findings strongly support the effectiveness of XR-enhanced MALL in improving contextual inference and accelerating lexical access among adult L2 learners. Initially, the study was carried out in a single intensive English program located in an urban setting where English is the primary language, giving participants authentic exposure to the second language outside of the classroom. While improving external validity for ESL contexts, this ecological advantage restricts the results' applicability to EFL contexts, where learners may need to rely more on controlled practice and explicit instruction because they do not regularly interact with the target language. To make up for the lack of ambient linguistic input in these settings, the cognitive scaffolding provided by XR might need to be more directive or pedagogically framed. Second, the intervention's four 45-minute sessions spread over two weeks were adequate to identify short-term processing effects, but not long-term retention, productive language use, or the durability of behavioral change. Lei et al. [15] caution that without systematic review and spaced reinforcement over long periods of time, short-term gains in MALL interventions frequently do not translate into long-lasting learning outcomes. Third, although the study adjusted for baseline proficiency and working memory, it did not account for individual differences in spatial ability or prior immersive technology experience. These factors may interact with the cognitive benefits of XR. For example, learners with lower spatial aptitude may encounter more needless cognitive load in 3D environments, counteracting the advantages of referential grounding. The following is a list of some significant research directions. Longitudinal research is needed to determine whether XR-mediated increases in comprehension efficiency result in steady improvements in academic literacy, speaking, or writing. Comparative studies across different linguistic and educational contexts are also necessary to ascertain whether the benefits of XR are consistent across varying degrees of linguistic distance and institutional support, especially in formal EFL classrooms in Asia, the Middle East, or Latin America. Future research should therefore concentrate on adaptive XR designs that modify immersion intensity according to learner profiles (e.g., spatial ability, L1 background, and proficiency level). This will surpass universally applicable adaptive cognitive scaffolding. These enhancements would align MALL with embodied cognition principles and the practical realities of global language education, where equity and accessibility must remain top design priorities.

7 CONCLUSION

This study shows that it is possible to successfully develop consistent study habits among health sciences students, especially during the critical first year, by combining project-based learning with a mobile app-enhanced, data-informed approach. Students were inspired to study more often and were able to understand the value of consistency when daily app-based practice was incorporated into a cooperative group project. Individual effort was converted into group practice through the team discussions and shared goals, and the mobile app provided objective evidence of improvement. Students moved from intermittent, reactive studying to a more planned, contemplative, and sustained habit, which resulted in more than just more time spent studying alone. Academic achievement improved in measurable ways as a result of this model's consistency. It is a powerful illustration of the self-directed, life-long learning that is expected of health professionals after they graduate because of its simplicity and reality. The framework provides enough structure and social support to promote the early development of these habits. Additionally, it offers a practical and scalable approach to rethink how professional training may help students thrive right away.

8 DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

Microsoft Copilot was the only program we used to proofread the English text. The authors are responsible for the content and accuracy of the final version.

9 REFERENCES

[1] A. Al-Abri, F. Ranjbaran Madiseh, and M. Morady Moghaddam, "Exploring learning-oriented language assessment in enhancing students' lexical fluency through MALL," *The Asia-Pacific Educ. Res.*, vol. 34, pp. 1–13, 2024. https://doi.org/10.1007/s40299-024-00832-7

[2] K. Almudibry, "A meta-analysis of the literature on mobile-assisted language learning in response to COVID-19 in Saudi Arabia," *World J. Engl. Lang.*, vol. 12, no. 8, pp. 106–112, 2022. https://doi.org/10.5430/wjel.v12n8p106

[3] I. Barth, E. Spector-Cohen, R. Sitman, G. Jiang, F. Liu, and Y. Xu, "Beyond small chunks," *Int. J. Comput. Assist. Lang. Learn. Teach.*, vol. 9, pp. 79–97, 2019. https://doi.org/10.4018/IJCALLT.2019040105

[4] G. G. Botero, F. Questier, and C. Zhu, "Self-directed language learning in a mobile-assisted, out-of-class context: Do students walk the talk?" *Comput. Assist. Lang. Learn.*, vol. 32, pp. 71–97, 2018. https://doi.org/10.1080/09588221.2018.1485707

[5] F. Cakmak, "Mobile learning and mobile-assisted language learning in focus," *Lang. Technol.*, vol. 1, no. 1, pp. 30–48, 2019. https://dergipark.org.tr/en/pub/lantec/issue/42816/517381

[6] V. Chan, "Using a virtual reality mobile application for interpreting learning: Listening to the students' voice," *Interact. Learn. Environ.*, vol. 32, no. 6, pp. 2438–2451, 2024. https://doi.org/10.1080/10494820.2022.2147958

[7] N. P. Daly, "Investigating learner autonomy and vocabulary learning efficiency with MALL," *Lang. Learn. Technol.*, vol. 26, no. 1, pp. 1–30, 2022. https://doi.org/10.64152/10125/73469

[8] R. Dashtestani and S. Hojatpanah, "Mobile-assisted language learning in a secondary school in Iran: Discrepancy between the stakeholders' needs and the status quo," in *Handbook for Online Learning Contexts: Digital, Mobile and Open: Policy and Practice*, Springer, 2021, pp. 157–174. https://doi.org/10.1007/978-3-030-67349-9_12

[9] S. Ebadi and A. Raygan, "Investigating the facilitating conditions, perceived ease of use and usefulness of mobile-assisted language learning," *Smart Learn. Environ.*, vol. 10, no. 1, pp. 1–15, 2023. https://doi.org/10.1186/s40561-023-00250-0

[10] R. Fithriani, "The utilisation of mobile-assisted gamification for vocabulary learning: Its efficacy and perceived benefits," *CALL-EJ*, vol. 22, no. 3, pp. 146–163, 2021.

[11] I. García-Martínez, J. M. Fernández-Batanero, D. Cobos Sanchiz, and A. Luque de la Rosa, "Using mobile devices for improving learning outcomes and teachers' professionalisation," *Sustainability*, vol. 11, no. 24, p. 6917, 2019. https://doi.org/10.3390/su11246917

[12] N. Ghorbani and S. Ebadi, "Exploring learners' grammatical development in mobile-assisted language learning," *Cogent Educ.*, vol. 7, no. 1, p. 1704599, 2020. https://doi.org/10.1080/2331186X.2019.1704599

[13] T. Gonulal, "The use of Instagram as a mobile-assisted language learning tool," *Contemp. Educ. Technol.*, vol. 10, no. 3, pp. 309–323, 2019. https://doi.org/10.30935/cet.590108

[14] M. F. Hafour, "The effects of MALL training on preservice and in-service EFL teachers' perceptions and use of mobile technology," *ReCALL*, vol. 34, no. 3, pp. 274–290, 2022. https://doi.org/10.1017/S0958344022000015

[15] X. Lei, J. Fathi, S. Noorbakhsh, and M. Rahimi, "The impact of mobile-assisted language learning on English as a foreign language learners' vocabulary learning attitudes and self-regulatory capacity," *Front. Psychol.*, vol. 13, p. 872922, 2022. https://doi.org/10.3389/fpsyg.2022.872922

[16] R. Li, "Effects of mobile-assisted language learning on EFL/ESL reading comprehension," *Educ. Technol. Soc.*, vol. 25, no. 3, pp. 15–29, 2022. https://www.j-ets.net/ETS/journals/25_3/2.pdf

[17] D. Li, N. Tang, M. Chandler, and E. Nanni, "An optimal approach for predicting cognitive performance in education based on deep learning," *Comput. Hum. Behav.*, vol. 167, p. 108607, 2025. https://doi.org/10.1016/j.chb.2025.108607

[18] M. Menekse *et al.*, "Enhancing student reflections with natural language processing-based scaffolding: A quasi-experimental study in a large lecture course," *Comput. Educ.: Artif. Intell.*, vol. 8, p. 100397, 2025. https://doi.org/10.1016/j.caeai.2025.100397

[19] J. Park, J. Paxtle-Granjeno, M. W. Ok, M. Shin, and E. Wilson, "Preventing digital distraction in secondary classrooms: A quasi-experimental study," *Comput. Educ.*, vol. 227, p. 105223, 2025. https://doi.org/10.1016/j.compedu.2024.105223

[20] S. Pan, B. Hafez, A. Iskandar, and Z. Ming, "Integrating constructivist principles in an adaptive hybrid learning system for developing social entrepreneurship education among college students," *Learn. Motiv.*, vol. 87, p. 102023, 2024. https://doi.org/10.1016/j.lmot.2024.102023

[21] D. R. Isbell, H. Rawal, R. Oh, and S. Loewen, "Narrative perspectives on self-directed foreign language learning in a computer- and mobile-assisted language learning context," *Languages*, vol. 2, no. 2, p. 4, 2017. https://doi.org/10.3390/languages2020004

[22] L. Hsu, "Examining EFL teachers' technological pedagogical content knowledge and the adoption of mobile-assisted language learning: A partial least square approach," *Comput. Assist. Lang. Learn.*, vol. 29, no. 8, pp. 1287–1297, 2016. https://doi.org/10.1080/09588221.2016.1278024

[23] C. Troussas, A. Krouska, and M. Virvou, "Integrating an adjusted conversational agent into a mobile-assisted language learning application," in *Proc. IEEE 29th Int. Conf. Tools Artif. Intell. (ICTAI)*, 2017, pp. 1153–1157. https://doi.org/10.1109/ICTAI.2017.00179

[24] S. Tarighat and S. Khodabakhsh, "Mobile-assisted language assessment: Assessing speaking," *Comput. Hum. Behav.*, vol. 64, pp. 409–413, 2016. https://doi.org/10.1016/j.chb.2016.07.015

[25] I. Fronza and D. Gallo, "Towards mobile-assisted language learning based on computational thinking," in *Int. Conf. Web-Based Learn*, Cham: Springer, 2016, pp. 141–150. https://doi.org/10.1007/978-3-319-47440-3_16

[26] R. Shadiev, W. Y. Hwang, and Y. M. Huang, "Review of research on mobile language learning in authentic environments," *Comput. Assist. Lang. Learn*, vol. 30, nos. 3–4, pp. 284–303, 2017. https://doi.org/10.1080/09588221.2017.1308383

[27] H. Sudo, Y. Noborimoto, and J. Takahashi, "Grit, self-efficacy, and study habits in mobile learning among physical therapy students," *International Journal for Educational Media and Technology*, vol. 19, no. 1, 2025.

[28] B. Moreillon *et al.*, "Prediction of plasma volume and total hemoglobin mass with machine learning," *Physiological Reports*, vol. 11, no. 19, p. e15834, 2023. https://doi.org/10.14814/phy2.15834

10 AUTHORS

Antony Desilva D. is a research scholar at the Department of English at B.S. Abdur Rahman Crescent Institute of Science and Technology, India. His research area is Neurolinguistic Programming.

Dr. Vijayakumar Selvaraj is an Associate Professor in the Department of English at B.S. Abdur Rahman Crescent Institute of Science and Technology, India, with a PhD in Applied Linguistics. His research is anchored in computer-assisted language learning, AI-driven pedagogical innovation, and inclusive digital education. Over 18 publications have appeared in Q1 journals such as The International Review of Research in Open and Distributed Learning, Humanities & Social Sciences Communications, and Transactions on Emerging Telecommunications Technologies. His interdisciplinary work integrates natural language processing, neuro-linguistic programming, and adaptive learning systems to enhance L2 proficiency, accessibility, and equity in digital learning environments. Several patents have been filed on AI-based educational tools, including assistive technologies for visually impaired learners. He supervises six doctoral candidates and serves as a peer reviewer for multiple Q1 journals. His scholarly output reflects a sustained commitment to evidence-based interventions that bridge pedagogy, technology, and linguistic equity (E-mail: vijayakumar@crescent.education).

Dr. Sathikulameen A. is an Assistant Professor and Research Supervisor in the Postgraduate & Research Department of English at The New College, Chennai. With 16 years of experience in education, he specializes in English Language Teaching, Language Acquisition and Pedagogy, Multimedia and Technology-Enhanced Language Learning, and Cultural Studies. He has published 18 research papers in various indexed journals, contributed to conference proceedings, and authored several books and chapters in edited volumes. He has also been actively involved in organizing and coordinating academic programs, workshops, and conferences. He has served as a resource person, keynote speaker, and examiner in numerous academic and professional development events.

Emmanuel Rajkumar B. is a research scholar at the Postgraduate and Research Department of English at The New College, University of Madras. His research areas include Ecocriticism, Cultural Studies, and English Language Teaching.