

Mobile-Assisted Language Learning Flow Features: A Systematic Review

Li Zi Jia

Faculty of Management Universiti Teknologi Malaysia, Johor Bahru, Malaysia
Email: lizijia@ieee.org

Noorminshah A. Iahad

Department of Applied Computing & Artificial Intelligence, Faculty of Computing
Universiti Teknologi Malaysia, Johor Bahru, Malaysia
Email: minshah@utm.my

DOI Link: <http://dx.doi.org/10.6007/IJARBS/v16-i9/28911>

Published Date: 13 September 2026

Abstract

Mobile-assisted language learning (MALL) extends practice beyond scheduled lessons, but the app-related conditions associated with flow remain unclear. This systematic review examined non-experimental cross-sectional quantitative studies and distinguished MALL features from learners' evaluations, task conditions, and learner or contextual factors. Searches of Scopus and Web of Science identified 89 records; after duplicate removal and screening, eight studies published between 2021 and 2026 met the eligibility criteria, representing 3,684 participants. Methodological quality was appraised using the 2024 CASP checklist for descriptive and cross-sectional studies, and findings were synthesised narratively because applications, flow measures, and analytical models varied. Playability showed the most consistent positive relationship with flow. Interactivity had no significant independent path in two studies, although it appeared in one sufficient high-flow configuration. Evidence for feedback was mixed, while collaboration or sharing was positively associated with flow in one study. Perceived usefulness and ease of use were generally associated with flow but represented learner evaluations rather than objective app features. Clear goals and challenge-skill balance were recurrent task-level correlates, and flow was associated with continuance intention, perceived learning, or satisfaction. Because the evidence was cross-sectional and predominantly self-reported, the findings indicate associations rather than causal feature effects.

Keywords: Mobile-Assisted Language Learning, Flow Experience, App Features, Technology Perceptions, Systematic Review

Introduction

Mobile-Assisted Language Learning

Mobile-assisted language learning (MALL) has developed from portable content delivery toward interaction, collaboration and learner-directed activity across settings (Kukulska-Hulme & Shield, 2008). Smartphones and language-learning applications allow learners to access instructional content in formal and informal contexts, practise in short sessions, and exercise control over learning time, place, pace and resources. Reviews of the field show continued growth in mobile vocabulary learning, gamification and application-based language practice (Duman et al., 2015; Ishaq et al., 2021). More recent bibliometric evidence also identifies learner autonomy, informal learning, artificial intelligence-supported personalisation and higher education as prominent areas (Karakaya & Bozkurt, 2022).

The presence of a mobile device does not itself explain educational value. Meta-analyses have reported positive average effects of mobile technologies on language learning, while also showing that outcomes vary across activities, delivery modes, duration and study quality (Sung et al., 2015; Mihaylova et al., 2022). A systematic review of real-language attainment reached a similar conclusion: effectiveness depends on the population, context and implementation of the technology (Figueiredo, 2023). Evidence from children's language and literacy learning further indicates that application features should be examined individually because different features do not produce uniform benefits (Booton et al., 2023). The relevant question is therefore not only whether a mobile application is used, but how it structures the learning experience.

Flow Experience

Flow refers to an optimal state of deep concentration and involvement in which an activity is experienced as intrinsically rewarding and action proceeds with a sense of control (Csikszentmihalyi, 1990). Flow is especially relevant to language learning because sustained attention and repeated practice are required, whereas mobile learning often occurs in fragmented environments with competing distractions. In the foreign-language classroom, flow has been characterised by challenge-skill balance, interest, control and focused attention (Egbert, 2003).

The transfer of flow theory to technology-mediated learning requires conceptual care. Reviews of computer-mediated environments have found considerable variation in how flow, its antecedents and its consequences are defined and measured (Finneran & Zhang, 2005; Valinatajbahnamiri & Siahtiri, 2021). A recent review of 123 empirical studies in second-language learning reached a similar conclusion: flow research is concentrated in Asian EFL and computer-mediated settings, commonly relies on componential Likert scales, and has examined task conditions and individual differences more often than change over time or objective learning outcomes (Aubrey et al., 2025). Flow should therefore be treated as an experiential state rather than used as a broad label for every positive response to technology.

The Relationship Between MALL App-Related Factors and Flow

Studies of MALL and flow examine interactivity, playability, feedback, collaboration, perceived usefulness, perceived ease of use, telepresence, motivation, clear goals, control and challenge-skill balance. These variables do not belong to one conceptual category. This review defines MALL features as technological, interface, social or instructional attributes

embedded in an application and open to designer intervention. Playability, interactive mechanisms, feedback systems and collaborative tools fit this definition. Perceived usefulness and ease of use are feature-related evaluations central to the Technology Acceptance Model rather than features themselves (Davis, 1989). Telepresence is a technology-use experience. Challenge-skill balance are task conditions, whereas motivation and mindset are learner variables. Satisfaction, perceived learning and continuance intention are treated as outcomes.

An association between a learner perception and flow does not show that a particular feature produced the experience. Adaptive difficulty, for example, can be coded as an app feature, whereas a learner's perceived challenge-skill balance is a task-related or psychological condition. An immediate corrective-feedback system belongs to the application design; a learner's rating of that feedback is an evaluation. Telepresence is likewise an experience of using a technology, not a built-in feature. Treating these constructs as interchangeable would obscure what the app provides and how learners respond to it.

The studies also indicate that these attributes may matter in combination. Interactivity and feedback can be non-significant in net-effect models yet appear within combinations of conditions associated with high flow (Meng & Li, 2026). Digital game-based vocabulary studies have linked playability, clear goals, feedback and challenge-skill balance with different dimensions of flow (Li et al., 2021; Wang & Feng, 2025), while collaborative MALL research has connected collaboration and sharing with flow and further adoption (Hu et al., 2023). Perceived usefulness and ease of use have also been associated with flow, satisfaction or continuance intention across app-based language-learning contexts (Lu et al., 2022; Wang et al., 2022). Aubrey et al. (2025) provide a broad methodological map of empirical research on flow across L2 learning, but their review was not designed to isolate modifiable MALL/app-related features or to distinguish those features systematically from learners' evaluations, task conditions and contextual factors. The existing review literature has therefore not yet synthesised which modifiable app-related attributes have been examined in relation to flow, how consistently they are associated with flow, or where the design-relevant evidence remains limited.

Current Review

Responding to this gap, the present systematic review examines non-experimental cross-sectional quantitative evidence on MALL and flow at the feature level. It advances the broader methodological synthesis of Aubrey et al. (2025) by focusing specifically on potentially modifiable MALL/app-related features and their reported relationships with flow. To avoid attributing flow to app design when the evidence concerns learners' responses or task characteristics, it separates MALL features from feature-related perceptions, task conditions, learner or contextual factors and outcomes. The review also compares the operationalisation of flow and appraises study quality so that its interpretations remain proportionate to the evidence. It addresses the following research questions:

RQ1: Which MALL features have been examined in relation to learners' flow experience, and what patterns of association have been reported?

RQ2: How are feature-related perceptions, task conditions, and learner or contextual factors associated with flow in MALL environments?

RQ3: What outcomes are associated with flow, and what methodological limitations constrain the interpretation of the existing evidence?

Methods

The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement (PRISMA 2020) (Page et al., 2021). The process comprised database searching, duplicate removal, title and abstract screening, full-text eligibility assessment, data extraction, narrative synthesis and methodological quality appraisal.

Search Strategy

The search was run in Scopus and Web of Science on 8 August 2026. Terms were organised into two blocks, one for flow and one for mobile or technology-supported language learning. A concise representation was: ("flow experience" OR "flow state" OR "learning flow") AND ("mobile-assisted language learning" OR "mobile language learning" OR "language learning app*" OR "digital game-based language learning" OR "technology-enhanced language learning"). The operational searches also included terms for immersion, concentration, gamification, artificial intelligence, virtual reality, augmented reality and vocabulary learning. Scopus used the full Boolean expression. In Web of Science, terms were searched with TS= and the language-learning and technology blocks were linked with NEAR/5. Feature terms were deliberately omitted; relevant feature evidence was coded from full texts retrieved by the broader MALL-flow search. Study design was also not used as a search filter because design labels are indexed inconsistently. Cross-sectional quantitative studies were selected during screening. Records from the two databases were merged, and duplicates were matched by DOI and title.

Eligibility Criteria

Studies were eligible when they reported original, non-experimental cross-sectional quantitative evidence; investigated second- or foreign-language learning in a mobile or app-based environment; explicitly measured flow; examined at least one app feature, system characteristic, technology-related perception or contextual condition in relation to flow; and provided sufficient information for extraction and appraisal. Experimental, quasi-experimental, longitudinal, qualitative and genuine mixed-methods studies were excluded. Meng and Li (2026) described their analysis as mixed methods because it combined SEM and fsQCA. Both components nevertheless analysed the same cross-sectional questionnaire dataset, and no independent qualitative data were reported. The study was therefore retained as cross-sectional quantitative evidence. Table 1 presents the full criteria.

Table 1

Inclusion and Exclusion Criteria

Criterion	Inclusion criteria	Exclusion criteria
Study type	Empirical, non-experimental cross-sectional quantitative studies using survey, correlational, comparative, mediation, SEM, network, profile, or configurational analysis of cross-sectional numerical data.	Qualitative, genuine mixed-methods, longitudinal, experimental, or quasi-experimental studies; reviews, meta-analyses, conceptual papers, editorials, and commentaries.
Learning context	English, EFL, ESL, or another clearly identified second- or foreign-language learning context.	Studies unrelated to second- or foreign-language learning.
Technology	Mobile language-learning apps, MALL systems, or app-based language-learning environments.	Studies without a mobile or app-based language-learning component.
Flow experience	Flow was explicitly measured or examined as a predictor, mediator, or outcome.	Studies discussing engagement, enjoyment, or immersion without explicitly measuring flow.
Research focus	At least one app feature, system characteristic, technology-related perception, task condition, or learner/contextual factor was examined in relation to flow.	Flow was not analysed in relation to the MALL environment or conditions of its use.
Language and access	English-language full text with sufficient information for extraction and appraisal.	Full text unavailable in English or insufficient methodological information.

Data Extraction

The review author used a standardised form to record publication details, educational context, sample, target language, application or platform, research design, flow measurement, predictors, outcomes, analytical method, principal findings and limitations. Variables were first extracted using the terminology of each study and then coded into five prespecified analytical categories: (a) MALL features or system characteristics, (b) feature-related perceptions or technology-use experiences, (c) task conditions, (d) learner-related or contextual factors, and (e) outcomes associated with flow. This two-stage procedure prevented motivation, mindset, perceived usefulness and other non-feature constructs from being relabelled as app features.

Quality Assessment

The review author assessed methodological quality with the 2024 Critical Appraisal Skills Programme checklist for descriptive and cross-sectional studies (Critical Appraisal Skills Programme, 2024). The same Q1-Q11 criteria were applied to all eight studies, with responses of Yes, No or Can't tell. Counts were retained for transparency. The meaning of each item is provided in the note to Table 4. No total score or categorical quality grade was calculated. Instead, the item-level judgements were interpreted across recruitment, measurement quality, sample adequacy, analytical rigour, risk of bias and applicability.

Data Synthesis

A narrative synthesis was conducted because the studies differed in application type, definitions of app-related attributes, flow instruments, populations and statistical models. The direction, magnitude where reported and statistical significance of relationships with flow were compared. Evidence was organised around app-related attributes, feature-related perceptions, task conditions, learner and contextual factors, and outcomes associated with flow. A meta-analysis was not undertaken because the constructs and model specifications were not sufficiently comparable to yield a meaningful pooled effect.

Results

Study Selection and Characteristics

The searches identified 89 records: 35 from Scopus and 54 from Web of Science. After 20 duplicates were removed by DOI and title, 69 records remained for title and abstract screening; 57 were excluded. Twelve reports were sought for retrieval, and one could not be obtained. Eleven full texts were assessed, of which three were excluded: one did not concern second- or foreign-language learning, one did not examine a MALL context, and one used a quasi-experimental design. Eight non-experimental cross-sectional quantitative studies were included. Figure 1 shows the selection process.

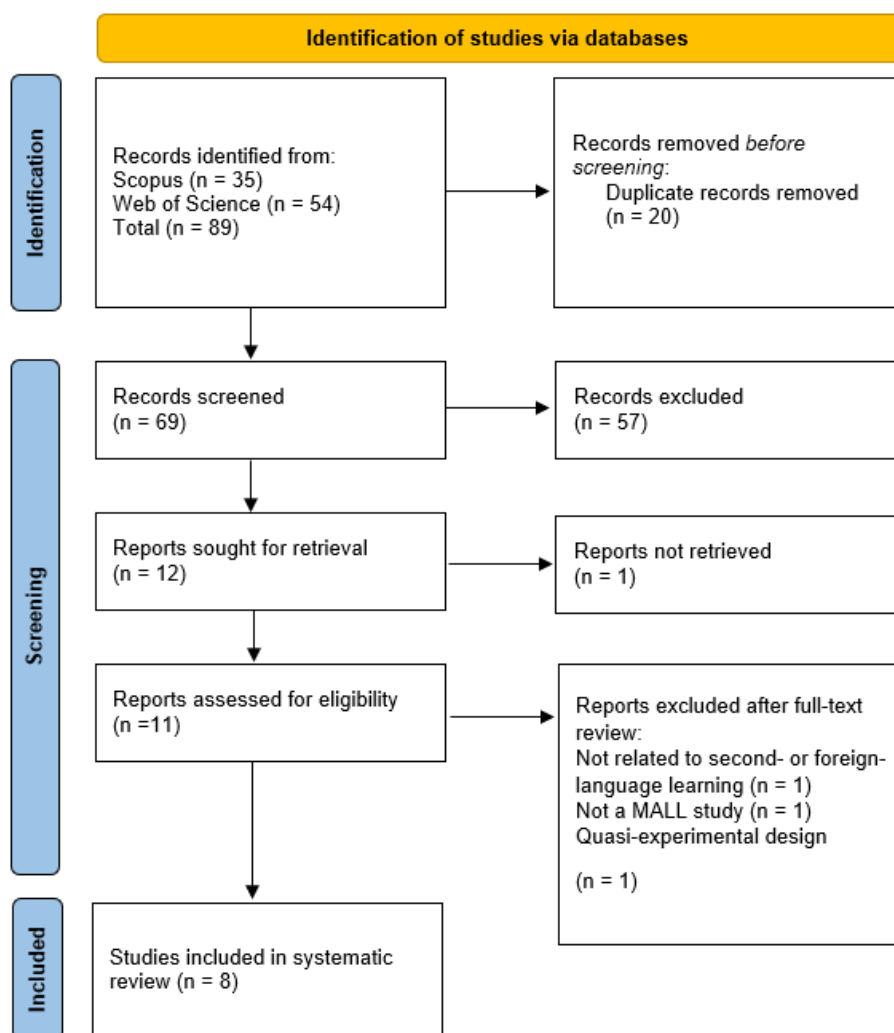


Figure 1. PRISMA flow diagram of the search and study-selection process

Therefore characterising the included evidence, it is important to distinguish the geographic concentration of the analytic sample from the scope of the wider field. Selected studies in South Korea and Iran have examined flow in smartphone-assisted listening and digitally gamified EFL learning. Table 2 presents these studies for geographical context only. They were not included in the eight-study synthesis because they fell outside the review's publication-window or cross-sectional-design criteria (Huh, 2016; Foroutan Far & Taghizadeh, 2024; Karimi & Nasouri, 2024).

Table 2

Selected International MALL-Flow Studies Outside the Analytic Sample

Study and context	Technology and flow focus	Reason outside analytic sample
Huh (2016) (Huh, 2016) South Korea; n = 30	Smartphone-app listening; flow, listening, and self-directed learning	Published before the 2021–2026 review window
Foroutan Far and Taghizadeh (2024) (Foroutan Far & Taghizadeh, 2024) Iran; n = 75	Digital gamification; collocation learning and sense of flow	Pre/post group intervention with qualitative components
Karimi and Nasouri (2024) (Karimi & Nasouri, 2024) Iran; n = 57	Digital text-based game; flow and incidental vocabulary learning	Randomised game comparison using diaries and outcome tests

Note. These studies illustrate the broader geographical reach of the field; they were not part of the eight-study synthesis and do not alter the PRISMA counts.

Although related MALL–flow research has been conducted in other countries, the eight studies meeting the present review's prespecified cross-sectional eligibility criteria were geographically concentrated in China. This concentration limits the cross-cultural generalisability of the synthesis but does not indicate that the wider field is confined to China. The eight studies meeting the prespecified eligibility criteria were published between 2021 and 2026 and represented a combined analytical sample of 3,684 participants; samples ranged from 289 to 664. The analytic evidence base was geographically concentrated: all eight studies were conducted in China or with learners studying in China, and most participants were university EFL learners. As Table 2 shows, related MALL–flow research has also been conducted in other national contexts, but those studies fell outside the present review's publication-window or design criteria. Seven studies used conventional cross-sectional survey designs analysed through structural equation modelling, mediation analysis, or regression. Meng and Li (2026) used a cross-sectional configurational design that combined PLS-SEM with fsQCA; it therefore remained quantitative and non-experimental. The environments included general English-learning apps, LAIX, Baicizhan, collaborative MALL, and digital game-based vocabulary learning. Five studies represented flow as a single or composite construct, two modelled concentration, intrinsic motivation, and enjoyment, and one used a nine-component flow score. Other analytical methods included regression-based mediation, latent profile analysis, and network analysis. Table 3 summarises these characteristics.

Table 3

Characteristics of the Included Studies

Study	Sample and context	MALL environment	Flow operationalisation	Analysis
Hu et al. (2023)	607 Chinese college students	Collaborative MALL	Single latent flow construct	SEM
Meng and Li (2026)	503 Chinese university EFL learners	Baicizhan competitive vocabulary learning	Overall flow-experience score	SEM and fsQCA
Wei and Li (2021)	289 Chinese university students	LAIX English-learning app	Single latent flow-experience construct	SEM
Li et al. (2021)	291 Chinese university EFL learners	Baicizhan vocabulary app	Concentration, intrinsic motivation and enjoyment	CFA and SEM
Lu, Liu and Chen (2025)	524 Chinese university students	English L2-learning apps	Nine-component flow score	Serial mediation and network analysis
Wang et al. (2022)	500 Chinese college students	English-learning apps	Composite of enjoyment, control and concentration	Regression and serial mediation
Lu et al. (2022)	664 CSL and ESL app users in China	Chinese- and English-learning apps	Composite learning-flow score	Latent profile and moderated regression
Wang and Feng (2025)	306 Chinese university EFL learners	Digital game-based vocabulary learning	Concentration, intrinsic motivation and enjoyment	SEM and multigroup analysis

Note. CFA = confirmatory factor analysis; CSL = Chinese as a second language; EFL = English as a foreign language; ESL = English as a second language; fsQCA = fuzzy-set qualitative comparative analysis; SEM = structural equation modelling.

Quality Assessment

The studies generally reported focused questions, established or adapted instruments, reliability or construct-validity evidence, and appropriate multivariate analyses. Recruitment was the most recurrent weakness: five studies received No and three received Can't tell for Q3. Sample-size adequacy was unclear in four studies because formal power calculations were not reported. Applicability to the wider MALL population was rated Can't tell in all eight studies because samples were restricted to particular universities, regions or learner groups. The appraisal was not converted into an overall quality category; interpretation focused on these recurring item-level risks and on their implications for the synthesis. Table 4 presents the item-level judgements.

Table 4

CASP Appraisal of the Eight Quantitative Cross-Sectional Studies

Study	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Hu et al. (2023)	Y	Y	CT	Y	Y	Y	Y	Y	Y	CT	Y
Meng and Li (2026)	Y	Y	N	Y	Y	Y	Y	Y	Y	CT	Y
Wei and Li (2021)	Y	Y	N	Y	Y	CT	Y	Y	Y	CT	Y
Li et al. (2021)	Y	Y	CT	Y	Y	CT	Y	Y	Y	CT	Y
Lu, Liu and Chen (2025)	Y	Y	N	Y	Y	Y	Y	Y	Y	CT	Y
Wang et al. (2022)	Y	Y	N	Y	Y	Y	Y	Y	Y	CT	Y
Lu et al. (2022)	Y	Y	CT	Y	Y	CT	Y	Y	Y	CT	Y
Wang and Feng (2025)	Y	Y	N	Y	Y	CT	Y	Y	Y	CT	Y

Note. Y = Yes; N = No; CT = Can't tell. Q1 = focused research question; Q2 = appropriate cross-sectional design; Q3 = acceptable recruitment; Q4 = accurate measurement with bias minimised; Q5 = data collection appropriate to the research question; Q6 = adequate sample size; Q7 = clear presentation of results; Q8 = sufficiently rigorous analysis; Q9 = clear statement of findings; Q10 = applicability to the review population; Q11 = value of the research (Critical Appraisal Skills Programme, 2024).

MALL Features and Flow*Interactivity and Playability*

Interactivity had no significant independent path to flow in either study that estimated one. Wei and Li (2021) reported $\beta = .061$ ($p = .353$), and Meng and Li (2026) reported $\beta = -.015$ ($p = .646$). In the latter study's fsQCA, interactivity nevertheless appeared in one sufficient high-flow configuration together with feedback, control and challenge-skill balance. This pattern distinguishes a non-significant net association from possible configurational relevance.

Playability showed the clearest positive feature-level pattern. Li et al. (2021) associated playability with concentration ($\beta = .486$, $p < .001$), but not intrinsic motivation or enjoyment. Wang and Feng (2025) reported positive paths to concentration ($b = .194$, $p = .013$), intrinsic motivation ($b = .227$, $p = .003$) and enjoyment ($b = .219$, $p = .005$). Both studies measured playability through learner reports rather than manipulating a discrete application component.

Feedback and Collaboration

Feedback produced mixed results. Li et al. (2021) reported a positive path to intrinsic motivation ($\beta = .126$, $p = .029$), but not to concentration or enjoyment. Wang and Feng (2025) reported positive paths to concentration ($b = .167$, $p = .016$), intrinsic motivation ($b = .211$, $p = .002$) and enjoyment ($b = .137$, $p = .047$). In Meng and Li (2026), feedback had no significant net association with overall flow ($\beta = .013$, $p = .713$), yet it was a core condition in all three

high-flow configuration types. Collaboration and sharing were positively associated with flow in one collaborative MALL study (Hu et al., 2023) ($\beta = .781, p < .001$).

Feature-Related Perceptions and Technology-Use Experiences

Perceived ease of use was positively associated with flow in three studies, although the pattern depended on context. Wei and Li (2021) reported $\beta = .417 (p < .001)$, and Meng and Li (2026) reported $\beta = .183 (p < .001)$. Lu et al. (2022) found a positive association for English-learning app users ($b = .15, p < .001$) but not Chinese-learning app users ($b = -.02, p = .44$). Perceived usefulness positively predicted flow in Wang et al. (2022) ($\beta = .319, p < .001$) and Lu, Liu and Chen (2025) ($\beta = .215, p < .001$); Lu et al. (2022) again found a positive association only in the English-learning group ($b = .14, p < .001$; CSL: $b = -.03, p = .21$). Telepresence was positively associated with flow in Wei and Li (2021) ($\beta = .295, p = .002$). Hu et al. (2023) reported a positive path from perceived cost to flow (path coefficient = .224) but no significant paths from perceived mobility or social influence. Accordingly, these variables were analysed as learner evaluations or technology-use experiences, not as objective properties of an app.

Task, Learner and Contextual Conditions

Clear goals and challenge-skill balance were recurrent correlates of flow. Meng and Li (2026) reported positive paths from clear goals ($\beta = .173, p < .001$), perceived control ($\beta = .466, p < .001$) and challenge-skill balance ($\beta = .234, p < .001$) to overall flow. Li et al. (2021) linked clear goals to concentration and enjoyment, and challenge-skill balance to concentration. Wang and Feng (2025) reported positive paths from both clear goals and challenge-skill balance to concentration, intrinsic motivation and enjoyment.

Learner and contextual variables also entered the models, but their direction must be distinguished from feature effects. In Wang et al. (2022), integrative motivation positively predicted flow ($\beta = .158, p < .001$) within the pathway perceived usefulness to integrative motivation to flow to continuance intention. App-use frequency was positively associated with flow in Wang et al. (2022) ($\beta = .157, p < .001$) and Lu, Liu and Chen (2025) ($\beta = .155, p < .001$). In the latter study, flow predicted motivation intensity ($\beta = .390, p < .001$) and language mindset ($\beta = .325, p < .001$); language mindset was therefore an outcome, not an antecedent, in that model. Lu et al. (2022) showed that target language and learning context moderated the relationships between usefulness, ease of use and flow.

Outcomes Associated With Flow

Flow was positively associated with continuance intention in Hu et al. (2023) ($\beta = .918, p < .001$), Wei and Li (2021) ($\beta = .371, p < .001$) and Wang et al. (2022) ($\beta = .204, p < .001$). In Wang et al. (2022), the indirect pathway from perceived usefulness through flow to continuance intention was .065 (95% CI [.032, .102]); the serial pathway through integrative motivation and flow was .006 (95% CI [.001, .014]). In digital game-based vocabulary learning, flow dimensions were generally associated with perceived learning and satisfaction. Li et al. (2021) found no significant enjoyment-satisfaction path ($\beta = .101, p = .125$), whereas Wang and Feng (2025) reported positive paths from all three flow dimensions to both outcomes. These are cross-sectional associations and indirect effects, not evidence of temporal mediation.

Table 5

Cross-Study Synthesis of Factors Examined in Relation to Flow

Factor	Analytical category	k	Relationship with flow	Synthesis
Playability	MALL Features (self-reported)	2	Positive in both studies; broader across flow dimensions in one study	Most consistent feature-level pattern
Interactivity	MALL Features (self-reported)	2	No significant independent path; present in one sufficient configuration	Configuration-dependent
Feedback	MALL Features (self-reported)	3	Mixed direct paths; present in all high-flow configurations in one study	Conditional rather than uniformly positive
Collaboration and sharing	MALL social feature (self-reported)	1	Positive association with flow	Promising but single-study evidence
Perceived ease of use	Feature-related perception	3	Generally positive; one study found an ESL but not CSL association	Context-dependent positive pattern
Perceived usefulness	Feature-related perception	3	Positive direct or indirect association with flow	Consistent perceptual correlate, not an objective feature
Perceived cost	Feature-related perception	1	Positive association with flow	Single-study evidence; interpretation depends on scale direction
Telepresence	Technology-use experience	1	Positive association with flow	Single-study evidence
Clear goals	Task condition	3	Positive overall or dimension-specific associations	Recurrent task-level correlate
Challenge-skill balance	Task-learner fit	3	Positive overall or dimension-specific associations	Recurrent task-level correlate
Perceived control	Learner experience of task	1	Positive association with overall flow	Limited evidence

Factor	Analytical category	k	Relationship with flow	Synthesis
Integrative motivation	Learner factor	1	Positive association with flow in a serial mediation model	Direction requires longitudinal confirmation
App-use frequency	Learner/contextual factor	2	Positive covariate association with flow	Observed, not experimentally assigned
Target language/context	Contextual factor	1	Moderated usefulness/ease-of-use relationships with flow	Flow relationships varied by learning context
Continuance intention	Outcome	3	Positive association with flow	Most recurrent downstream correlate
Perceived learning/satisfaction	Outcomes	2	Generally positive; one enjoyment-satisfaction path was non-significant	Limited to game-based vocabulary models
Motivation intensity/language mindset	Outcomes	1	Flow positively predicted both variables	Direction tested in one cross-sectional model

note. k = number of included studies examining the factor in relation to flow. Counts describe this eight-study evidence base and do not represent pooled effect sizes.

Discussion

Summary of Principal Findings

The eight studies collectively indicate that flow in MALL is associated with a combination of application attributes, learner perceptions, task conditions and contextual factors rather than with a single app feature. Playability showed the most consistent feature-level pattern, whereas feedback and interactivity were conditional on the flow dimension or configuration examined. Perceived usefulness, perceived ease of use and telepresence were generally associated with flow, but they represented learner evaluations or technology-use experiences. Clear goals and challenge-skill balance were recurrent task-level correlates. Flow was also positively associated with continuance intention and, in several models, with perceived learning or satisfaction. These findings answer the three research questions while retaining the conceptual and methodological limits of cross-sectional evidence.

MALL Features and Flow

Across the studies, no single feature explained flow. Playability showed the clearest positive pattern, especially for concentration, but both studies asked learners to rate playability rather than manipulating game components. The evidence therefore cannot determine whether competition, scoring, time pressure, visual presentation or another element produced the association. Variation in the design and evaluation of gamified MALL has also been noted in

reviews (Ishaq et al., 2021). At present, playability is best treated as a candidate design condition rather than a cause of flow.

Interactivity was not independently associated with flow in the two direct-path models. Its appearance within a sufficient configuration in Meng and Li (2026) suggests that interaction may matter when combined with feedback, learner control and an appropriate challenge-skill balance. Thus, the quality and task relevance of interaction may be more important than its quantity. This interpretation is consistent with broader reviews warning that flow in computer-mediated environments depends on the alignment of person, artefact and task rather than on technological stimulation alone (Finneran & Zhang, 2005; Valinatajbahnamiri & Siahtiri, 2021).

Feedback and collaboration were more conditional. Feedback had different direct relationships across the three studies, while its presence in every sufficient configuration reported by Meng and Li (2026) suggests that it may operate jointly with ease of use, goals, control or challenge-skill balance. Collaboration and sharing showed a strong association with flow in one study, but a single estimate cannot be generalised across applications. Feedback and social tools may be relevant when they reduce uncertainty or create purposeful participation; their mere presence is not evidence that they will produce flow.

Perceptions, Task Conditions and Context

Perceived usefulness, perceived ease of use and telepresence were associated with flow in several studies, but they remain learner evaluations or experiences. Their association with flow may reflect design quality, prior expectations, successful task completion or common response tendencies. The difference between Chinese- and English-learning groups, together with non-significant results for perceived mobility and social influence, shows that favourable technology perceptions do not operate uniformly. App features may shape perceptions, but the present data cannot treat perceptions as direct proxies for objective design.

Clear goals and challenge-skill balance were among the most recurrent correlates of flow. Perceived control was examined less often but showed a positive association in the one study that modelled it. These conditions correspond to foundational accounts of flow in language learning (Egbert, 2003) and to dimensions frequently identified in the recent review of L2 flow research (Aubrey et al., 2025). The pattern suggests that application design should be evaluated by how it structures the task, not by the number of functions it contains.

Integrative motivation, app-use frequency and target-language context further indicate that flow is situated. Importantly, variables changed roles across models. Integrative motivation preceded flow in Wang et al. (2022), whereas flow preceded motivation intensity and language mindset in Lu, Liu and Chen (2025). Combining these variables under a generic label of 'learner antecedents' would reverse the direction of evidence in the latter study. The defensible interpretation is that flow reflects an alignment among application, task, learner and context, with temporal direction still unresolved.

Theoretical Contributions

This review makes three bounded theoretical contributions. First, it provides a category system that distinguishes MALL features from feature-related perceptions, technology-use

experiences, task conditions, learner or contextual factors and outcomes. This distinction connects MALL design research with flow theory without treating perceived usefulness and ease of use, which originate in technology-acceptance research, as objective application features (Davis, 1989). It also retains clear goals, perceived control and challenge-skill balance as conditions of the learning activity and experience, consistent with established accounts of flow in language learning (Egbert, 2003; Aubrey et al., 2025).

Second, the evidence suggests a relational rather than feature-deterministic account of MALL flow. Interactivity and feedback did not operate uniformly as independent predictors, but both appeared within configurations that combined application, task and learner conditions. This pattern extends person-artefact-task accounts of flow in computer-mediated environments by showing why the presence of a technical affordance may be insufficient when its task relevance and learner experience are not considered (Finneran & Zhang, 2005; Valinatajbahnamiri & Siahtiri, 2021).

A final contribution concerns model direction. Wang et al. (2022) placed integrative motivation before flow, but Lu, Liu and Chen (2025) placed flow before motivation intensity and language mindset. Combining all three variables under the label of flow antecedents would therefore misrepresent the models that were actually tested. Since the data were cross-sectional, these directions are treated as modelling choices rather than evidence of temporal order.

Practical Implications

For designers, the evidence favours coherence between an app and its learning task over a larger function count. Playability was linked most consistently to concentration, although the studies did not isolate particular game mechanics. Feedback may be useful when learners can interpret it and connect it to a clear goal. Interactivity and collaboration also need a defined language-learning purpose, because interactivity alone had no independent association with flow. This emphasis is consistent with broader MALL reviews showing that learning outcomes vary with activity design, delivery mode, implementation and the features of the application (Ishaq et al., 2021; Sung et al., 2015; Mihaylova et al., 2022; Figueiredo, 2023). These remain evidence-informed design considerations rather than experimentally established effects.

Design evaluation should record objective provisions and learner perceptions separately. For example, whether an application provides adaptive difficulty, corrective feedback or collaboration is distinct from whether learners experience the application as useful, easy to use or socially engaging. Feature-level reporting would help developers identify whether weak flow is associated with the absence of a component, its implementation or learners' interpretation of it. This recommendation is consistent with evidence that different mobile application features do not produce uniform language and literacy outcomes (Booton et al., 2023).

For teachers and programme designers, app selection should be accompanied by task design. Goals can be made explicit, activities can be sequenced so that challenge develops gradually, and learners can be supported in interpreting feedback. Teachers should also monitor whether interface effort, unfamiliar game rules or collaborative requirements divert attention from language practice. The contextual differences reported by Lu et al. (2022) indicate that

an app associated with flow in one target-language setting may not operate similarly in another. Practical decisions should therefore combine evidence about the application with evidence about the learner group, task and instructional context.

Limitations and Future Research Directions

Evidence remains limited. The eight studies were cross-sectional, relied mainly on convenience or self-selected university samples, and measured the key constructs by self-report. Moreover, although related international studies exist (Table 2), the analytic sample was concentrated in Chinese higher education, leaving uncertain how the findings apply to school pupils, community-based adult learners, other target languages or different sociocultural settings. The frequent absence of formal power calculations, together with CASP concerns about recruitment and applicability, further limits confidence.

Coverage of the review also has limits. The search included Scopus and Web of Science but not specialist databases such as ERIC or PsycINFO. Because the search addressed MALL and flow broadly rather than requiring feature terms, studies using other terminology or indexed only elsewhere may have been missed. Only retrievable English-language reports were eligible. Experimental, quasi-experimental, longitudinal and qualitative studies were excluded to preserve a cross-sectional evidence base, which narrowed the forms of evidence available. Screening, coding and appraisal were conducted by one reviewer, so selection and classification decisions were not independently checked. The review protocol was not prospectively registered.

Heterogeneity in applications, predictor definitions, flow measures and analytical specifications prevented a meaningful meta-analysis. Standardised coefficients from different models were therefore reported descriptively and were not compared as if they were common effect sizes. Publication bias could not be assessed reliably with eight diverse studies. The categories in Table 5 should consequently be read as a map of the available evidence, not as a ranking of causal feature effectiveness; causal claims and recommendations require particular caution when they are derived from non-intervention research (Brady et al., 2023).

Future research should test the temporal and causal pathways suggested by this synthesis. Longitudinal studies could establish whether changes in app perceptions and flow precede later continuance behaviour. Experiments and quasi-experiments could manipulate adaptive difficulty, feedback timing, learner control, playability or collaborative interaction. Factorial and component-level designs would help separate individual features from the configurations in which they operate. Comparability would also benefit from broader samples, preregistered analyses, reported power calculations and more consistent flow measures.

Future studies should combine questionnaires with behavioural records, such as app-use logs, task completion, response time and persistence, and should measure flow repeatedly during use. Reports should also distinguish what an app objectively provides from how learners rate usefulness, ease of use, feedback or telepresence. With those distinctions, future reviews could ask a sharper question: which modifiable attributes are related to flow, for which learners, and in which learning conditions?

Conclusion

With respect to RQ1, this review identified four potentially modifiable app-related attributes examined in relation to flow: playability, interactivity, feedback, and collaboration or sharing. Playability showed the most consistent positive feature-level relationship, particularly with concentration. Evidence for feedback varied by flow dimension or configuration, interactivity had no significant independent relationship with flow, and collaboration or sharing was positively related to flow in one study. Although the playability finding was based on only two studies using learner-reported measures, its consistency makes playability the clearest candidate for component-level experimental testing. It should not, however, be interpreted as evidence that gamification or any particular game mechanic causes flow.

With respect to RQ2, the synthesis showed that several frequently studied variables were not app features. Perceived usefulness and ease of use were learner evaluations, telepresence was a technology-use experience, and clear goals, perceived control and challenge-skill balance represented task or experiential conditions. Distinguishing these variables from modifiable app attributes prevents favourable learner evaluations from being treated as evidence that a particular design feature produced flow. The available evidence instead points to an alignment among app design, task structure, learners' perceptions and the learning context.

With respect to RQ3, flow was positively associated with continuance intention, perceived learning and satisfaction in several models, but the methodological evidence substantially constrains their interpretation. All included studies were cross-sectional, most constructs were measured through self-report, and the evidence was concentrated in Chinese higher education. The review itself was also limited to Scopus and Web of Science, included only retrievable English-language reports, and relied on a single reviewer for screening, coding and quality appraisal. Taken together, the review contributes a feature-oriented evidence map showing which MALL attributes may be modifiable, which variables instead represent learner or task conditions, and where the evidence remains insufficient. The findings identify associations and priorities for further testing rather than causal feature effects. Future longitudinal and experimental studies should manipulate clearly defined app attributes—especially specific components underlying playability—record objective app use and examine flow over time across more diverse learner groups and target languages.

References

- Aubrey, S., Zhang, L., & Zhou, M. R. (2025). Flow in language learning: A systematic review of methodological features. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688251389358>
- Booton, S. A., Hodgkiss, A., & Murphy, V. A. (2023). The impact of mobile application features on children's language and literacy learning: A systematic review. *Computer Assisted Language Learning*, 36(3), 400–429. <https://doi.org/10.1080/09588221.2021.1930057>
- Brady, A. C., Griffin, M. M., Lewis, A. R., Fong, C. J., & Robinson, D. H. (2023). How scientific is educational psychology research? The increasing trend of squeezing causality and recommendations from non-intervention studies. *Educational Psychology Review*, 35, Article 37. <https://doi.org/10.1007/s10648-023-09759-9>
- Critical Appraisal Skills Programme. (2024). *CASP checklist: For descriptive/cross-sectional studies*. <https://casp-uk.net/casp-tools-checklists/>
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Duman, G., Orhon, G., & Gedik, N. (2015). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, 27(2), 197–216. <https://doi.org/10.1017/S0958344014000287>
- Egbert, J. (2003). A study of flow theory in the foreign language classroom. *The Modern Language Journal*, 87(4), 499–518. <https://doi.org/10.1111/1540-4781.00204>
- Figueiredo, S. (2023). The effect of mobile-assisted learning in real language attainment: A systematic review. *Journal of Computer Assisted Learning*, 39(4), 1083–1102. <https://doi.org/10.1111/jcal.12811>
- Finneran, C. M., & Zhang, P. (2005). Flow in computer-mediated environments: Promises and challenges. *Communications of the Association for Information Systems*, 15, 82–101. <https://doi.org/10.17705/1CAIS.01504>
- Foroutan Far, F., & Taghizadeh, M. (2024). Comparing the effects of digital and non-digital gamification on EFL learners' collocation knowledge, perceptions, and sense of flow. *Computer Assisted Language Learning*, 37(7), 2083–2115. <https://doi.org/10.1080/09588221.2022.2146724>
- Hu, L., Hei, D., Wang, H., & Dai, X. (2023). Chinese college students collaborative mobile-assisted language learning experience and flow as a key factor for further adoption. *Frontiers in Psychology*, 14, Article 1165332. <https://doi.org/10.3389/fpsyg.2023.1165332>
- Huh, K. (2016). The effect of flow experience on English listening and self-directed learning abilities through a listening activity using a smartphone application. *Multimedia-Assisted Language Learning*, 19(3), 158–177. <https://doi.org/10.15702/mall.2016.19.3.158>
- Ishaq, K., Mat Zin, N. A., Rosdi, F., Jehanghir, M., Ishaq, S., & Abid, A. (2021). Mobile-assisted and gamification-based language learning: A systematic literature review. *PeerJ Computer Science*, 7, Article e496. <https://doi.org/10.7717/peerj-cs.496>
- Karakaya, K., & Bozkurt, A. (2022). Mobile-assisted language learning (MALL) research trends and patterns through bibliometric analysis: Empowering language learners through ubiquitous educational technologies. *System*, 110, Article 102925. <https://doi.org/10.1016/j.system.2022.102925>

- Karimi, M. N., & Nasouri, A. (2024). EFL learners' flow experience and incidental vocabulary learning during text-based game tasks: The moderating role of working memory capacity. *System*, 124, Article 103398. <https://doi.org/10.1016/j.system.2024.103398>
- Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(3), 271–289. <https://doi.org/10.1017/S0958344008000335>
- Li, R., Meng, Z., Tian, M., Zhang, Z., & Xiao, W. (2021). Modelling Chinese EFL learners' flow experiences in digital game-based vocabulary learning: The roles of learner and contextual factors. *Computer Assisted Language Learning*, 34(4), 483–505. <https://doi.org/10.1080/09588221.2019.1619585>
- Lu, A., Deng, R., Huang, Y., Song, T., Shen, Y., Fan, Z., & Zhang, J. (2022). The roles of mobile app perceived usefulness and perceived ease of use in app-based Chinese and English learning flow and satisfaction. *Education and Information Technologies*, 27(7), 10349–10370. <https://doi.org/10.1007/s10639-022-11036-1>
- Lu, A., Liu, S., & Chen, W. (2025). Perceived usefulness of English (L2) learning apps and language mindset mediated by flow and motivation intensity: A serial mediation model and a network analysis. *Interactive Learning Environments*, 33(2), 1735–1754. <https://doi.org/10.1080/10494820.2024.2375637>
- Meng, J., & Li, R. (2026). Configurational antecedents of flow experience in EFL vocabulary acquisition: A mixed-methods analysis. *International Journal of Applied Linguistics*. Advance online publication. <https://doi.org/10.1111/ijal.70231>
- Mihaylova, M., Gorin, S., Reber, T. P., & Rothen, N. (2022). A meta-analysis on mobile-assisted language learning applications: Benefits and risks. *Psychologica Belgica*, 62(1), 252–271. <https://doi.org/10.5334/pb.1146>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, Article 105906. <https://doi.org/10.1016/j.ijsu.2021.105906>
- Sung, Y.-T., Chang, K.-E., & Yang, J.-M. (2015). How effective are mobile devices for language learning? A meta-analysis. *Educational Research Review*, 16, 68–84. <https://doi.org/10.1016/j.edurev.2015.09.001>
- Valinatajbahnamiri, M., & Siahtiri, V. (2021). Flow in computer-mediated environments: A systematic literature review. *International Journal of Consumer Studies*, 45(4), 511–545. <https://doi.org/10.1111/ijcs.12640>
- Wang, X., & Feng, L. (2025). Examining the influential mechanism of English as a foreign language (EFL) learners' flow experiences in digital game-based vocabulary learning: Shedding new light on a priori proposed model. *Education Sciences*, 15(2), Article 125. <https://doi.org/10.3390/educsci15020125>
- Wang, X., Lu, A., Lin, T., Liu, S., Song, T., Huang, X., & Jiang, L. (2022). Perceived usefulness predicts second language learners' continuance intention toward language learning applications: A serial multiple mediation model of integrative motivation and flow. *Education and Information Technologies*, 27(4), 5033–5049. <https://doi.org/10.1007/s10639-021-10822-7>

Wei, N., & Li, Z. (2021). Telepresence and interactivity in mobile learning system: Its relation with open innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, 7(1), Article 78. <https://doi.org/10.3390/joitmc7010078>