

Bridging Gaps with Technology: A Systematic Review of Digital Scaffolding in ESL Classrooms

Afza Diyana Abdullah, Nurul Ashikin Izhar* & Mohd Ali Samsudin

Pusat Pengajian Ilmu Pendidikan, Universiti Sains Malaysia, 11700 Gelugor, Pulau Pinang, Malaysia

*Corresponding author email: ashikinizhar@usm.my

ARTICLE HISTORY

Received: 18 August 2025

Revised: 16 November 2025

Accepted: 15 December 2025

Publisher: 20 January 2026

KEYWORDS

digital scaffolding

ESL

systematic literature review

educational technology

blended learning

ABSTRACT - Digital scaffolding has gained increasing attention in English as a Second Language (ESL) education; however, its conceptualisation and implementation remain fragmented. This systematic literature review synthesises empirical research on digital scaffolding in ESL contexts, examining how it is defined, the factors driving its adoption, and how it is implemented across formal educational settings. Guided by PRISMA 2020 protocols, a structured Scopus search identified 214 records published between 2020 and 2025, of which 8 studies met the inclusion criteria after screening. The synthesis reveals persistent conceptual ambiguity: while some studies equate digital scaffolding with technological affordances such as automated feedback or digital prompts, others frame it as adaptive, contingent, and fading support grounded in sociocultural learning theory. Drivers of digital scaffolding are broadly shared but contextually framed, with global studies emphasising enrichment, learner autonomy, and personalization, and Malaysian studies highlighting equity-related constraints such as feedback scarcity, infrastructural challenges, and assessment demands. Implementation spans micro-level tools, meso-level instructional sequences, and macro-level learning ecosystems, yet most interventions remain skill-specific, short-term, and limited in attention to scaffold fading and transfer. Overall, the review demonstrates that digital scaffolding extends beyond technology use and requires pedagogically designed, layered support that is sensitive to learner development and contextual conditions.

INTRODUCTION

Scaffolding is a foundational construct in educational theory, originating from sociocultural perspectives on learning and the concept of the Zone of Proximal Development, which emphasises the role of guided support in enabling learners to perform beyond their current level (Vygotsky, 1978). In language education, scaffolding has traditionally referred to temporary and responsive assistance provided by teachers or peers, such as modelling, questioning, prompting, and feedback (Wood et al., 1976; Walqui, 2006). A defining feature of scaffolding is that such support is gradually withdrawn as learners gain competence and independence (Wood et al., 1976).

With the increasing integration of digital technologies into English as a Second Language (ESL) instruction, scaffolding has extended beyond face-to-face interaction to include digitally mediated forms of support. Technology-enhanced learning environments now structure tasks, guide attention, and provide feedback that supports learners' engagement with complex content (Reiser, 2004). As a result, the term *digital scaffolding* has emerged to describe instructional support that is delivered through or facilitated by digital means while retaining its pedagogical function.

Despite its growing use, digital scaffolding remains inconsistently conceptualised in ESL research. In some accounts, scaffolding is treated as a pedagogical process that preserves its core characteristics of intentionality, contingency, and fading, albeit mediated through technology (Walqui, 2006; Reiser, 2004).

In other cases, the term is used more loosely to describe the presence of digital tools or resources, without clear articulation of how such supports function instructionally or whether they are designed to be withdrawn. This conceptual slippage blurs the distinction between scaffolding as a pedagogical construct and technology as a delivery mechanism.

The need for digital scaffolding is often justified by persistent challenges faced by ESL learners in technology-mediated environments. Learners commonly experience cognitive overload, difficulties in organising language output, and limited self-regulation when instructional guidance is reduced (Reiser, 2004). From a sociocultural perspective, scaffolding addresses these challenges by structuring participation and guiding learners toward independent performance through mediated support (Vygotsky, 1978; Walqui, 2006).

Given these issues, there is a need for a focused synthesis that clarifies how digital scaffolding is defined and operationalised in ESL contexts. Rather than equating scaffolding with technology use, such a synthesis must examine whether instructional scaffolding principles articulated in sociocultural theory are retained when support is mediated through digital environments (Wood et al., 1976; Reiser, 2004). Addressing this gap is essential for strengthening conceptual clarity and informing principled instructional design in technology-enhanced ESL education.

PROBLEM STATEMENT

Despite sustained investment in educational technologies and digital pedagogies, ESL learners' speaking proficiency remains persistently underdeveloped, particularly in contexts where English is learned as a second or foreign language. Speaking continues to be reported as the most challenging language skill due to its real-time cognitive demands, affective barriers such as anxiety, and limited opportunities for meaningful oral interaction beyond the classroom (Adickalam & Md Yunus, 2022; J. Wang et al., 2022). Large-scale reviews and bibliometric analyses consistently show that, while reading and writing receive substantial instructional and research attention, speaking instruction remains comparatively under-theorised and under-supported, especially in technology-mediated environments (Wang et al., 2022; Zhangli et al., 2024).

The rapid shift toward technology-enhanced, blended, and online learning environments has further complicated the teaching and learning of speaking skills. Although digital tools such as mobile applications, learning management systems, video-based platforms, and AI-driven applications are increasingly adopted, empirical findings on their effectiveness for speaking development remain fragmented and inconsistent (Ramalingam et al., 2022; Zhangli et al., 2024). Several studies report positive outcomes in terms of motivation, engagement, or reduced speaking anxiety (Budianto et al., 2025; Ding & Muhyiddin, 2025), while others highlight challenges related to cognitive overload, superficial interaction, limited feedback quality, and uneven learner participation (Chen et al., 2023; Ma et al., 2022). This inconsistency suggests that technology alone does not guarantee effective speaking development.

A critical issue underlying these mixed findings is the lack of pedagogically grounded scaffolding in technology-mediated speaking instruction. While scaffolding is widely recognised as essential for supporting learners' progression from assisted to independent performance, many digital speaking interventions rely heavily on fixed prompts, task instructions, or automated feedback without sufficient adaptivity, contingency, or gradual fading of support (Hasan & Bidin, 2023; Wang et al., 2025). Reviews of blended and online ESL instruction indicate that scaffolding is often concentrated at pre-task or post-task stages, leaving the actual speaking performance phase under-supported or overly reliant on learners' self-regulation abilities (Ng et al., 2025; Ramalingam et al., 2022).

Moreover, existing systematic reviews in ESL research tend to aggregate speaking with other language skills or focus narrowly on specific technologies such as mobile learning, gamification, or flipped classrooms (Chuane et al., 2023; Kernagaran & Abdullah, 2022). As a result, there is limited synthesis that explicitly examines how digital scaffolding is conceptualised, operationalised, and evaluated for speaking development. Recent reviews on AI-assisted language learning and blended instruction acknowledge this gap, calling for more skill-specific, theory-informed syntheses that move beyond tool-centric descriptions (Li & Zhao, 2025; Zhangli et al., 2024).

Consequently, there is a clear need for a systematic literature review that critically examines digital scaffolding approaches for ESL speaking, focusing on how scaffolding principles are embedded within technology-enhanced environments, what speaking sub-skills are targeted, and what learning outcomes are reported. Addressing this gap is essential for informing more coherent pedagogical design, guiding future empirical research, and supporting practitioners in making principled decisions about technology use in ESL speaking instruction.

RESEARCH QUESTIONS

In light of these gaps, this review is guided by three questions:

RQ1: How is digital scaffolding defined and conceptualised in ESL research?

RQ2: What pedagogical and learning-related factors drive the implementation of digital scaffolding in ESL contexts?

RQ3: How is digital scaffolding implemented in ESL contexts?

These research questions are designed to address the conceptual and methodological fragmentation surrounding digital scaffolding in ESL research. By first examining how digital scaffolding is defined and theorised (RQ1), the review establishes a common conceptual baseline against which studies can be meaningfully compared. The second question (RQ2) situates digital scaffolding within the pedagogical and learning challenges it is intended to address, thereby clarifying the conditions under which scaffolding is deemed necessary rather than assuming its inherent value. The third question (RQ3) shifts the focus from definition to enactment, analysing how scaffolding is operationalised in digital environments through instructional design, mediated interaction, and support mechanisms. These questions enable a systematic synthesis that moves beyond descriptions of technology use to a principled understanding of digital scaffolding as a pedagogical construct in ESL contexts.

METHODOLOGY

Review Design and Protocol

This study adopted a systematic literature review (SLR) design to synthesise empirical research on digital scaffolding in ESL contexts. The review was conducted in accordance with the PRISMA 2020 guidelines (Page et. al., 2021) which provide a transparent and replicable framework for identifying, screening, and synthesising relevant studies. An SLR approach was deemed appropriate because the aim of the study was not to evaluate the effectiveness of specific technologies, but to critically examine how digital scaffolding has been conceptualised, justified, and implemented across empirical ESL research.

The review was guided by a concept-driven synthesis strategy, whereby digital scaffolding served as the primary analytical construct rather than a predefined set of tools or platforms. This approach allowed for close examination of how scaffolding principles were interpreted and operationalised in digital environments and reduced the risk of conflating technology use with pedagogical scaffolding. The review protocol was defined a priori, including eligibility criteria, search strategy, study selection procedures, quality appraisal, and data extraction methods, to minimise selection bias and enhance methodological rigour.

Eligibility Criteria

Clear inclusion and exclusion criteria were established to ensure that only studies directly relevant to the conceptual examination of digital scaffolding in ESL contexts were retained. The criteria were defined prior to the screening process and were guided by the review's focus on how digital scaffolding is defined, justified, and implemented.

Inclusion criteria were as follows:

- empirical studies situated in ESL or EFL contexts
- studies involving digital, online, blended, or technology-mediated learning environments
- studies that explicitly or implicitly described instructional support aligned with scaffolding principles, such as guided tasks, modelling, prompts, feedback, or staged assistance

- studies published in peer-reviewed journals
- studies published in English between 2020 and 2025, reflecting contemporary post-pandemic digital learning contexts

Exclusion criteria included:

- conceptual papers, opinion pieces, editorials, or book chapters without empirical data
- studies focused solely on technology acceptance, usability, or attitudes without pedagogical analysis of instructional support
- studies in which technology use was reported without evidence of instructional design or mediated support relevant to scaffolding
- non-ESL/EFL contexts or studies targeting languages other than English
- non-peer-reviewed publications such as theses, conference abstracts, or reports.

These criteria ensured that the final corpus comprised studies in which digital scaffolding could be examined as a pedagogical construct rather than inferred from general technology use.

Information Sources and Search Strategy

The literature search was conducted exclusively using Scopus. The search targeted empirical studies related to digital scaffolding and technology-mediated instructional support in Malaysian secondary ESL contexts. A comprehensive Boolean search string was constructed and applied to the TITLE, ABSTRACT, and KEYWORDS fields to ensure relevance while maintaining sufficient breadth.

To maximise coverage and avoid premature exclusion of relevant studies, the search strategy did not impose overly specific constraints on speaking skills or the explicit use of the term *digital scaffolding* at the database retrieval stage. Preliminary testing of highly targeted combinations resulted in extremely limited retrieval. This reflects common reporting practices in ESL research, where speaking outcomes are frequently embedded within broader communicative, interactional, or task-based constructs, and where scaffolding is often operationalised implicitly through guided tasks, feedback mechanisms, or instructional sequencing rather than labelled explicitly as “digital scaffolding.” Consequently, restricting the search to narrowly defined speaking terminology and explicit scaffolding labels at the identification stage would risk excluding substantively relevant studies. Following established systematic review practices, broader ESL and technology-enhanced learning descriptors were therefore prioritised during database searching, with speaking-related outcomes and scaffolding characteristics identified and verified during subsequent screening stages.

The following search string was applied on 19 August 2025:

```
(TITLE-ABS-KEY(("ESL") OR ("English as a Second Language") OR ("EFL") OR ("English as a Foreign Language") OR ("TESOL") OR ("English language learning") OR ("second language learning") OR ("foreign language learning")))  
AND (("digital scaffolding") OR ("technology enhanced language learning") OR ("technology-enhanced language learning") OR ("computer assisted language learning") OR ("CALL") OR ("mobile assisted language learning") OR ("MALL") OR ("technology enhanced learning") OR ("online language learning") OR ("blended learning") OR ("intelligent tutoring system") OR ("AI-assisted learning") OR ("AI tool") OR ("AI tools") OR ("artificial intelligence")))  
AND ((Secondary school*) OR (high school*))  
AND ("Malaysia"))  
AND (PUBYEAR > 2019 AND PUBYEAR < 2026)  
AND (LIMIT-TO (AFFILCOUNTRY,"Malaysia"))  
AND (LIMIT-TO (SUBJAREA,"SOCI") OR LIMIT-TO (SUBJAREA,"ARTS"))  
AND (LIMIT-TO (LANGUAGE,"English"))
```

This search yielded an initial total of 214 records, which were subsequently exported for screening and eligibility assessment in accordance with PRISMA guidelines.

Selection Process

The study selection process followed the PRISMA 2020 framework and proceeded through four stages: identification, screening, eligibility assessment, and inclusion. The flow of study selection is presented in Figure 1. At the identification stage, the Scopus database search using the predefined Boolean string yielded 214 records. During the export and reference management process, 18 duplicate records were identified and removed. The remaining 196 records were retained for screening.

During the screening stage, titles and abstracts of the 196 records were reviewed to exclude studies that were clearly outside the scope of the review. A total of 142 records were excluded at this stage. The primary reasons for exclusion included studies that were not situated in ESL/EFL contexts, did not involve school-level learners, focused on technology use without instructional or pedagogical support, or examined learner perceptions and technology acceptance without reference to instructional design. Following this stage, 54 records were retained for full-text assessment.

At the eligibility stage, full-text versions of the 54 remaining articles were assessed against the predefined inclusion criteria. Forty-six studies were excluded after full-text review. The majority of these exclusions were due to implicit or non-instructional uses of technology, where digital tools were described without explicit instructional scaffolding or structured learner support. Additional exclusions occurred due to contextual mismatches, including inappropriate instructional level or study settings that did not align with the review focus.

At the inclusion stage, a final set of 8 studies met all eligibility criteria and were included in the qualitative synthesis. These studies explicitly or partially operationalised digital scaffolding in ESL contexts and provided sufficient pedagogical detail to address at least one of the review's research questions.

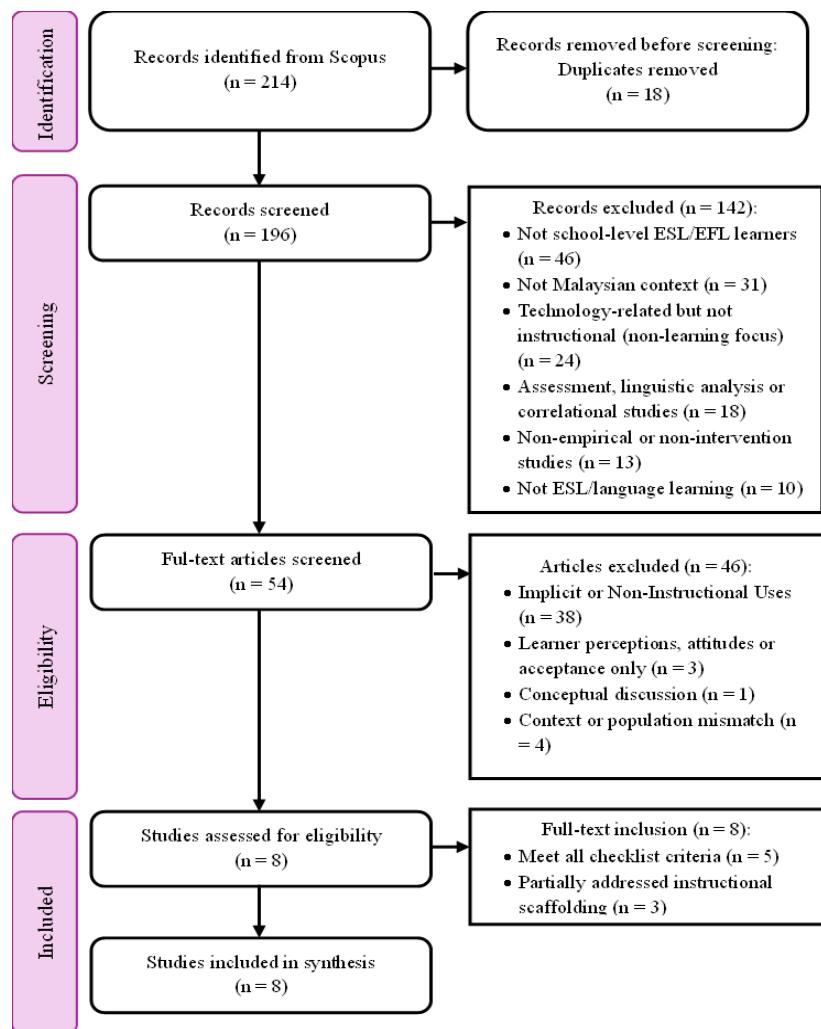


Figure 1. PRISMA flow diagram

Quality Appraisal

The methodological quality of the included studies was appraised using the Mixed Methods Appraisal Tool (MMAT), version 2018. The MMAT was selected because it allows for the systematic evaluation of qualitative, quantitative, and mixed-methods studies within a single review, which was appropriate given the methodological diversity of the included articles. The tool assesses studies based on five core criteria specific to each study design, focusing on clarity of research questions, appropriateness of data collection and analysis, coherence between data and interpretation, and consideration of methodological limitations.

Each of the eight included studies was first categorised according to its primary research design (qualitative, quantitative, or mixed methods) and then appraised against the corresponding MMAT criteria. To maintain transparency and avoid artificial precision, studies were not excluded solely on the basis of MMAT scores. Instead, quality appraisal findings were used to contextualise the strength and credibility of the evidence during synthesis, particularly when interpreting how digital scaffolding was conceptualised and implemented.

The methodological quality of the included studies was appraised using the Mixed Methods Appraisal Tool (Hong et al., 2018). As summarised in Table 1, the studies demonstrated overall moderate to good methodological quality, with recurring limitations related to short intervention duration, convenience sampling, and limited integration in mixed methods designs.

However, common methodological limitations were also observed, including limited justification of sampling strategies, short intervention durations, and insufficient discussion of researcher positivity or bias in qualitative designs. In several quantitative studies, outcome measures were reported without detailed validation procedures, which constrained the interpretation of causal claims.

These quality considerations informed the synthesis by foregrounding conceptual clarity and instructional transparency over effect size or outcome magnitude. As the purpose of this review was to examine definitions, drivers, and implementation of digital scaffolding rather than to determine effectiveness, all studies meeting the inclusion criteria were retained for analysis, with methodological limitations explicitly acknowledged.

Table 1. Methodological Quality Appraisal of Included Studies Using MMAT (Hong et al., 2018)

Study	Design Category (MMAT)	Clarity of Research Questions	Appropriateness of Design	Data Collection Adequacy	Sampling Strategy	Data Analysis Rigor	Integration of Data (if applicable)	Overall Appraisal
Chew et al. (2019)	Quantitative (Quasi-experimental)	Clear	Appropriate	Pre/post-tests clearly described	Convenience sampling	Appropriate statistical analysis	N/A	Moderate to Good
Hasan & Bidin (2023)	Qualitative	Clear	Aligned with pedagogical focus	Classroom artefacts & observations	Single-site purposive sample	Thematic analysis described	N/A	Moderate
Shukor, Chuane, Albakri, Madzlan & Gopal (2025)	Quantitative (Quasi-experimental)	Clear	Suitable for intervention study	Tests and questionnaires	Limited sample size	Statistical procedures reported	N/A	Moderate to Good
Erni et al. (2023)	Quantitative	Clear	Appropriate	Instruments described	Convenience sampling	Descriptive & inferential statistics	N/A	Moderate
Tinggie, Tan, Muslim & Keng (2023)	Qualitative	Clear	Appropriate	Observations, interviews, learner work	Small purposive sample	Systematic coding process	N/A	Moderate
Ng, Azlan, Kamal & Manion (2020)	Mixed Methods	Clear	Suitable mixed-methods design	Quantitative + qualitative data	Sampling not fully justified	Separate analyses conducted	Integration implicit	Moderate
Zhou, Lee & Kew (2025)	Quantitative	Clear	Appropriate	Performance measures reported	Context-specific sampling	Statistical analysis adequate	N/A	Moderate to Good
Nair, Zainudin, Krishnasamy, & Siddique (2025)	Mixed Methods	Clear	Appropriate	Surveys, tasks, reflections	Sampling constraints	Qualitative & quantitative analyses	Limited explicit integration	Moderate

Data Extraction and Synthesis

Data extraction was conducted using a structured extraction framework designed to align directly with the review's research questions. For each included study, key information was systematically recorded, including publication details, educational context, participant level, digital learning environment, research design, and reported learning focus. In addition, analytical categories were developed to capture how digital scaffolding was defined or conceptualised, the pedagogical or learning-related drivers for its use, and the instructional strategies through which it was implemented.

To address RQ1, data were extracted on how each study described or implied scaffolding, with particular attention to whether scaffolding was framed as a pedagogical strategy, a design feature embedded within digital environments, or an emergent form of support arising through interaction. For RQ2, extraction focused on the challenges or conditions that justified the use of scaffolding, such as linguistic difficulty, cognitive load, learner anxiety, limited autonomy, or reduced interaction in online or blended settings. For RQ3, data were extracted on specific instructional practices, including task sequencing, modelling, prompts, feedback mechanisms, peer-mediated support, and the extent to which support was faded or sustained over time.

A qualitative narrative synthesis approach was employed to integrate findings across studies. Rather than aggregating outcomes or effect sizes, the synthesis compared patterns in how scaffolding was conceptualised and operationalised across contexts. Studies were grouped analytically based on the explicitness of scaffolding, distinguishing between explicitly designed instructional scaffolding and partially or implicitly realised support. This categorisation enabled cross-study comparison while preserving methodological and contextual differences.

Throughout the synthesis process, methodological quality findings from the MMAT appraisal were used to contextualise interpretations, particularly where conceptual claims were weakly supported by empirical detail. The synthesis prioritised conceptual coherence and instructional transparency, enabling a critical examination of digital scaffolding as a pedagogical construct in ESL contexts rather than as a proxy for technology use.

Contextual Scope

Although the search strategy was designed to capture studies situated in secondary school ESL contexts in Malaysia, the final corpus of included studies spans both school-based and higher education settings. Of the eight studies reviewed, three were conducted in primary or secondary ESL classrooms (Tinggie et al., 2023; Shukor et al., 2025; Nair et al., 2025), while five were situated in higher education ESL/EFL environments (Chew et al., 2020; Ng et al., 2020; Hasan & Bidin, 2023; Erni et al., 2023; Zhou et al., 2024). This imbalance reflects a broader pattern in digital scaffolding research, where empirical investigations are more frequently undertaken in tertiary settings due to greater institutional support, technological readiness, and assumptions of learner autonomy. As such, the inclusion of higher education studies is not incidental but indicative of the current research landscape in technology-enhanced ESL learning.

While speaking skill development was identified as a primary area of interest, the review found a limited number of studies that explicitly examined digital scaffolding for speaking in isolation, particularly within school-based contexts. Preliminary searches that combined speaking-specific outcomes with explicit digital scaffolding constructs yielded a very small corpus. To avoid an overly restrictive scope and to ensure sufficient empirical coverage, the review therefore includes multi-skill ESL studies in which speaking is embedded within broader instructional designs, with speaking-related findings examined where relevant.

To address this cross-level and multi-skill scope, the review adopts a transferability-oriented interpretive stance rather than direct generalisation. Findings from higher education studies are analysed as sources of pedagogical and design principles that may inform secondary ESL practice, rather than as prescriptive models. Higher education research tends to conceptualise digital scaffolding as design-mediated and self-regulatory, embedded in task sequencing, feedback systems, and platform affordances (Chew et al., 2020; Ng et al., 2020; Zhou et al., 2024). In contrast, school-based studies emphasise scaffolding as compensatory and stabilising support, responding to mixed proficiency levels, limited instructional time, and the need for repeated guided practice (Tinggie et al., 2023; Shukor et al., 2025). Recognising these contextual distinctions is essential for interpreting the review findings, as it foregrounds how digital scaffolding operates differently across ESL settings while preserving secondary education as the primary analytical lens.

RESULTS & DISCUSSION

RQ1 Defining Digital Scaffolding in ESL

When synthesised across the eight included studies, digital scaffolding in ESL research emerges as a conceptually unstable construct, not because of overt theoretical disagreement, but because of uneven attention to scaffolding principles. Although all studies describe forms of learning support mediated by digital environments, they diverge in where scaffolding is assumed to reside (whether in instructional design, technological systems, or learner interaction) leading to different degrees of theoretical coherence (Chew et al., 2019; Ng et al., 2020; Tinggie et al., 2023).

Studies that conceptualise digital scaffolding as explicit instructional design demonstrate the strongest alignment with canonical scaffolding theory. Across these studies, scaffolding is framed as an intentional pedagogical strategy that structures learners' engagement with tasks through staged activities, modelling, and guided practice. For instance, Chew et al. (2019) and Ng et al. (2020) explicitly distinguish scaffolding from technology use, emphasising that learning gains depend on how instructional support is sequenced rather than on the digital tool itself. Similarly, Hasan and Bidin (2023) conceptualise scaffolding as teacher-led mediation, where digital tools function as supports for pedagogical intent rather than autonomous instructional agents. Taken together, these studies suggest that conceptual clarity is achieved when scaffolding is anchored in instructional decision-making, with technology serving a mediational role.

In contrast, several studies implicitly relocate scaffolding from pedagogy to design-mediated or tool-based support. In these cases, scaffolding is inferred from system features such as inquiry prompts, structured interfaces, or multimodal task layers, rather than articulated as an instructional strategy. Erni et al. (2023), for example, conceptualise scaffolding through strategy-based guidance embedded in an online reading environment, while Zhou et al. (2025) frame scaffolding as layered task design within a digital multiliteracies platform. Although these studies acknowledge learner challenges such as cognitive overload and engagement, scaffolding is primarily associated with environmental affordances rather than adaptive instructional mediation. Across the corpus, this shift corresponds with limited discussion of contingency or fading, indicating a conceptual drift toward viewing scaffolding as a static design feature.

A third conceptual orientation frames scaffolding as emergent or interactional support, particularly through peer-mediated learning. Tinggie et al. (2023) conceptualise scaffolding as peer assistance enacted through questioning and negotiation during writing tasks, drawing on sociocultural assumptions about collaborative learning. Similarly, Nair et al. (2025) describe scaffolding through learners guided interaction with digital graphic organisers, where support emerges from how learners use visual tools rather than from explicit instructional sequencing. While these studies demonstrate the social dimension of learning, scaffolding is often identified retrospectively and remains weakly specified in pedagogical terms. When examined collectively, these studies blur the boundary between scaffolding and collaborative learning, as instructional responsibility and fading mechanisms are rarely articulated.

Across all three conceptual orientations, a critical pattern is the systematic absence of fading as an explicit analytical concern. Even in studies that claim to implement scaffolding, support is typically sustained throughout the intervention without clear evidence of gradual withdrawal or transfer of responsibility to learners (Chew et al., 2019; Erni et al., 2023; Zhou et al., 2025). This suggests that digital scaffolding is frequently conceptualised as continuous assistance rather than temporary support, departing from canonical definitions and raising concerns about learner dependency.

The synthesis indicates that digital scaffolding in ESL research is best understood as a continuum of conceptualisations rather than a unified construct. Studies that foreground instructional design and pedagogical intent demonstrate greater theoretical coherence, while tool-centred and emergent conceptualisations risk diluting scaffolding into a descriptive label for technology use or interaction (Ng et al., 2020; Tinggie et al., 2023; Nair et al., 2025). This conceptual variability underscores the need for greater theoretical precision in defining digital scaffolding and cautions against equating digital environments with scaffolded instruction.

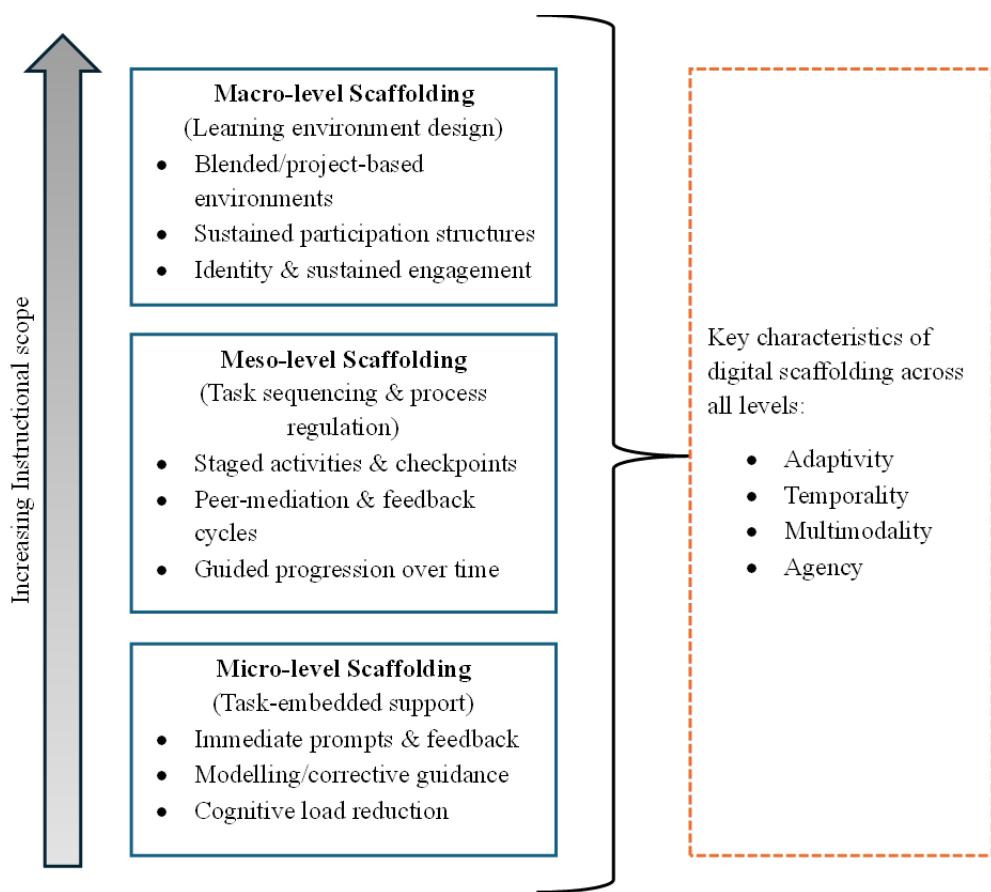


Figure 2. An Analytically Derived Model of Digital Scaffolding in ESL Contexts

These patterns suggest that digital scaffolding in ESL research is not confined to isolated instructional moves, but is distributed across multiple layers of instructional organisation. Building on this synthesis, the RQ1 findings were distilled into the layered analytical model shown in Figure 2, which captures where scaffolding is most consistently enacted across the reviewed studies. The model conceptualises digital scaffolding as operating at three interrelated levels: micro-level scaffolding embedded within moment-to-moment learner–task interaction, meso-level scaffolding realised through the sequencing and regulation of learning activities over time, and macro-level scaffolding embedded in the design of learning environments and participation structures. Across all three levels, scaffolding is shaped by recurring characteristics (adaptivity, temporality, multimodality, and agency) that influence how support is delivered, sustained, and transferred to learners. Figure 2 represents a synthesis grounded in the empirical patterns identified across the eight studies, illustrating digital scaffolding as a layered pedagogical construct rather than a single technological or instructional feature.

While RQ1 clarifies how digital scaffolding is conceptualised across ESL research, these conceptualisations also implicitly raise questions about why such scaffolding is deemed necessary. The layered model presented in Figure 2 suggests that scaffolding is invoked in response to challenges that emerge at different levels of instructional organisation: immediate cognitive and linguistic demands during task performance, difficulties in sustaining progression across learning activities, and structural constraints within broader learning environments. Accordingly, the need for digital scaffolding cannot be attributed to a single deficit or instructional gap, but rather to a constellation of learner-, task-, and context-related pressures. Building on this framework, RQ2 examines the pedagogical and contextual drivers that propel the use of digital scaffolding in ESL contexts, with particular attention to how these drivers align with the micro-, meso-, and macro-levels identified in the RQ1 synthesis.

RQ2 What Factors Propel the Need for Digital Scaffolding in ESL Contexts?

Across the eight reviewed studies, the need for digital scaffolding emerges not from isolated instructional shortcomings, but from recurrent learner and instructional needs that surface when ESL learning is mediated through digital or blended environments. Synthesised across studies, these needs cluster around four dominant patterns: the need for cognitive stabilisation during task engagement, the need for sustained instructional guidance beyond task initiation, the need for regulated participation and interaction, and the need for compensatory support in constrained learning environments. These needs recur across contexts and technologies, suggesting that digital scaffolding is invoked as a response to systemic misalignments between learner capacities, instructional demands, and learning conditions.

The most consistently reported need is the need for cognitive and linguistic stabilisation during task engagement. Several studies document learners' difficulty in managing idea organisation, comprehension, and language accuracy when engaging with digital tasks independently. Chew et al. (2019) show that learners struggle to summarise and structure ideas in online writing tasks without guided support, while Erni et al. (2023) report cognitive overload in online reading environments characterised by dense input and limited strategic awareness. Hasan and Bidin (2023) further demonstrate that lower-achieving learners are particularly vulnerable to performance breakdowns in the absence of explicit instructional guidance. These findings indicate that digital environments amplify learners' immediate processing demands, creating a need for scaffolding that anchors cognition and prevents disengagement at critical points of task execution.

A second recurring need concerns the absence of sustained guidance once tasks are initiated. Several studies reveal that learners' engagement deteriorates when digital tasks lack instructional continuity. Ng et al. (2020) demonstrate that unguided mobile learning results in fragmented participation and superficial task completion, while Nair et al. (2025) highlight learners' difficulty in retaining and organising information across writing stages. These studies suggest that scaffolding is driven not merely by entry-level difficulty, but by the need to maintain instructional momentum over time, particularly in environments where learners are expected to self-regulate.

A third pattern relates to the need to regulate participation and interaction in digitally mediated learning spaces. Tinggie et al. (2023) identify reduced peer interaction and over-reliance on teacher intervention in digital writing contexts, prompting the use of peer-mediated scaffolding to re-establish collaborative meaning-making. Zhou et al. (2024) similarly frame scaffolding as necessary to sustain learner engagement and agency within digital platforms. These findings indicate that digital environments do not automatically foster interaction; instead, scaffolding is often required to structure participation and prevent passive or uneven engagement.

A fourth and contextually salient need concerns structural and environmental constraints that limit learners' access to English input and practice. In Malaysian ESL contexts, several studies point to restricted instructional time, limited exposure outside the classroom, and logistical challenges associated with blended learning. Shukor et al. (2025) position scaffolding as necessary to support repetition, guided practice, and continuity within constrained instructional conditions. In this sense, scaffolding functions as a compensatory mechanism, addressing systemic limitations rather than learner deficits alone.

When these needs are considered in relation to the micro–meso–macro framework established in RQ1, an important critical insight emerges. The needs do not map neatly onto a single level of scaffolding; rather, they cut across levels, revealing tensions between how scaffolding is needed and how it is conceptualised. Cognitive stabilisation aligns most visibly with micro-level scaffolding, yet studies often rely on static or persistent supports that lack planned fading. The need for sustained guidance and regulated participation points toward meso-level scaffolding, but several studies implement support unevenly, addressing task sequencing without fully supporting learner self-regulation. Structural constraints motivate macro-level scaffolding, yet system-level designs sometimes substitute environmental access for pedagogical mediation, risking over-reliance on platforms rather than instructional intent. As summarised in Table 2, the reviewed studies consistently report needs related to cognitive stabilisation, instructional continuity, participation regulation, and structural constraints. Importantly, these needs do not map neatly onto single levels of scaffolding, revealing tensions between why scaffolding is required and how it is conceptualised across levels.

This misalignment suggests that while the needs for digital scaffolding are clearly articulated, the corresponding scaffolding responses are not always theoretically coherent across levels. Scaffolding is frequently introduced to address immediate problems (overload, disengagement, lack of interaction) without explicit consideration of how support should evolve across micro-, meso-, and macro-levels. As a result, scaffolding risks becoming permanent support rather than temporary mediation, particularly in contexts where digital tools are expected to compensate for broader instructional constraints.

The findings indicate that digital scaffolding is driven by persistent and interrelated pedagogical needs intensified by digital learning environments, but that these needs are not always matched by level-appropriate or developmentally responsive scaffolding designs. This critical tension underscores the importance of examining not only *why* scaffolding is needed, but also *how* it is subsequently implemented.

Table 2. Needs Driving Digital Scaffolding in ESL Contexts

Emergent Need Pattern	How the Need Manifests Across Studies	Illustrative Evidence from Reviewed Studies	Relationship to the Scaffolding Framework
Cognitive and linguistic stabilisation during task engagement	Learners struggle with idea organisation, comprehension, and language accuracy when engaging independently with digital tasks; cognitive overload is frequently reported	Online writing tasks requiring summarisation and organisation without guidance (Chew et al., 2019); dense online reading environments overwhelming learners' processing capacity (Erni et al., 2023); low-achieving learners unable to cope without explicit instructional support (Hasan & Bidin, 2023)	Aligns most strongly with micro-level scaffolding, yet studies often rely on persistent rather than fading supports, indicating incomplete enactment of scaffolding temporality
Sustained instructional guidance beyond task initiation	Learners disengage or complete tasks superficially once initial instructions are given; difficulty maintaining progression across stages of learning	Unguided mobile learning leading to fragmented engagement (Ng et al., 2020); learners unable to retain and organise ideas across writing phases (Nair et al., 2025)	Points to meso-level scaffolding, but highlights a gap between task sequencing and genuine support for self-regulation
Regulation of participation and interaction in digital spaces	Reduced peer interaction, uneven participation, or over-reliance on teacher input in digital learning contexts	Limited peer engagement in digital writing tasks (Tinggie et al., 2023); need to sustain learner agency and engagement in online platforms (Zhou et al., 2024)	Suggests meso- and macro-level needs, yet scaffolding responses are often interaction-enabling rather than developmentally contingent
Compensatory support for structural and environmental constraints	Limited exposure to English, reduced instructional time, and logistical constraints in blended or online contexts	Need for repeated, guided practice within constrained Malaysian secondary settings (Shukor et al., 2025); reliance on structured digital environments to offset limited contact time (Ng et al., 2020)	Motivates macro-level scaffolding, but risks substituting access-oriented design for pedagogical mediation

continued

Prevention of learner disengagement and breakdown (cross-cutting)	Learners withdraw cognitively or affectively when digital tasks exceed their independent capabilities	Disengagement observed when scaffolding is absent or poorly sequenced across studies (Chew et al., 2019; Ng et al., 2020; Tinggie et al., 2023)	Cuts across all levels, underscoring the need for coherent scaffolding trajectories rather than isolated supports
--	---	---	---

RQ3 Implementation of Digital Scaffolding in ESL

The reviewed studies indicates that digital scaffolding is implemented through a range of instructional strategies that vary in explicitness, scope, and responsiveness to learner needs. Rather than following a single implementation model, studies adopt different scaffolding mechanisms to address cognitive, procedural, interactional, and contextual challenges identified in RQ2. However, the synthesis also reveals uneven alignment between identified needs and implemented scaffolding designs, resulting in both effective mediation and conceptual slippage.

A prominent implementation pattern involves explicit scaffolding embedded directly within tasks, particularly to address learners' cognitive and linguistic difficulties. Chew et al. (2019) implement scaffolding through system-guided writing stages, modelling, and task constraints that guide learners step-by-step through complex writing processes. Hasan and Bidin (2023) similarly employ modelling, prompts, and staged drafting activities to support learners' descriptive writing, with digital tools serving as delivery mechanisms for teacher-led guidance. Erni et al. (2023) implement strategy-based scaffolding through inquiry prompts and guided reading strategies to regulate learners' comprehension processes in online reading environments. Across these studies, scaffolding is explicit, visible, and instructional, directly responding to the need for cognitive stabilisation identified in RQ2. However, while such implementations are effective in supporting task completion, they often rely on sustained support throughout the intervention, with limited attention to planned fading. This raises questions about whether scaffolding functions as temporary mediation or as permanent instructional support.

A second implementation pattern addresses learners' difficulty in sustaining engagement and progression across learning activities. Ng et al. (2020) implement scaffolding by structuring mobile learning tasks into guided sequences, demonstrating that learning outcomes improve only when activities are deliberately orchestrated rather than left unguided. Nair et al. (2025) employ digital graphic organisers to scaffold learners' planning, organisation, and revision across multiple writing stages, positioning scaffolding as a process that unfolds over time rather than as isolated support. These implementations align with the need for sustained instructional guidance identified in RQ2, yet they also expose a recurring limitation. While task sequencing is clearly articulated, learner self-regulation and transition to independence are rarely operationalised explicitly. Scaffolding is present across stages but mechanisms for reducing support or transferring control to learners are often implicit.

To address participation and interactional needs, some studies implement scaffolding through peer-mediated structures. Tinggie et al. (2023) operationalise scaffolding through peer questioning, elicitation, and collaborative problem-solving during digital writing tasks. Rather than relying on teacher or system control, scaffolding emerges through guided interaction among learners, supported by digital platforms. This approach directly targets the interactional challenges identified in RQ2, particularly reduced peer engagement in digital environments. However, peer-mediated scaffolding is highly contingent on learner competence and participation norms, making it uneven in quality and difficult to regulate. Without explicit instructional framing or monitoring, peer scaffolding risks becoming inconsistent or superficial, blurring the distinction between collaborative activity and pedagogically intentional scaffolding.

At a broader level, digital scaffolding is implemented through learning environment and curriculum design. Zhou et al. (2024) implement scaffolding by embedding layered task designs and multimodal prompts within a digital multiliteracies platform, distributing support across tasks rather than concentrating it in teacher feedback. Shukor et al. (2025) implement scaffolding at the system level by sequencing grammar tasks, repetition, and guided practice within a blended learning environment to compensate for limited instructional time in Malaysian secondary contexts. These implementations respond directly to the structural and contextual needs identified in RQ2. However, they also reveal a

critical tension: when scaffolding is embedded primarily at the environmental level, instructional mediation risks being replaced by access-oriented design. In such cases, scaffolding may support participation and exposure without necessarily ensuring contingent, learner-responsive support.

Digital scaffolding implementation in ESL contexts reflects a continuum from highly explicit instructional mediation to diffuse, environment-level support. Studies that implement scaffolding explicitly within tasks tend to preserve pedagogical intent but risk over-supporting learners. In contrast, studies that rely on system- or environment-level scaffolding promote flexibility and access but often under-specify instructional contingency and fading. Crucially, the review reveals that implementation frequently prioritises addressing immediate learning breakdowns over designing coherent scaffolding trajectories across time and levels. As a result, scaffolding is often implemented as a set of discrete strategies rather than as an integrated, developmentally responsive system.

Digital Scaffolding from ESL Perspective

From an ESL perspective, the reviewed studies do not conceptualise digital scaffolding around isolated language skills (e.g., speaking, writing, or reading), but rather around conditions that enable second language development in constrained instructional contexts. This finding aligns with the RQ1 synthesis, which showed that scaffolding is defined less as a discrete instructional technique and more as a layered orchestration of supports embedded across tasks, tools, and learning environments. Unlike EFL settings, where skill development is often compartmentalised and exposure is limited, ESL learning environments are characterised by uneven proficiency levels, limited classroom time, high curricular demands, and strong examination pressures (Ag-Ahmad, Mohamed & Majilang, 2025). Such conditions induce digital scaffolding not as skill-specific remediation but as a systemic pedagogical response to contextual constraints as it shapes how learners' access, engage with, and sustain participation in language tasks.

Across RQ2, a consistent pattern emerges: the need for digital scaffolding is driven by misalignment between task demands and learner readiness, rather than by linguistic deficiency alone. Several studies illustrate how learners struggle when cognitively or linguistically complex tasks are introduced without staged guidance. For instance, Chew et al. (2020) and Hasan and Bidin (2023) demonstrate that higher education learners encounter difficulty when academic literacy tasks are presented as complete performances rather than scaffolded processes. Similarly, school-based studies such as Tinggie et al. (2023) and Shukor et al. (2025) show how interactional and grammatical demands overwhelm learners when support is absent, delayed, or insufficiently structured. These findings suggest that scaffolding in ESL contexts functions as a bridging mechanism that mediate between curricular expectations and learners' developmental readiness in settings where instructional pacing often exceeds learner capacity.

RQ3 further reveals that, in practice, digital scaffolding is implemented predominantly through design-mediated and system-supported mechanisms. It positions technology as a compensatory tool rather than a transformational pedagogy. Across the reviewed studies, digital platforms are frequently used to substitute for instructional resources that are difficult to sustain in face-to-face classrooms, such as continuous feedback, extended practice time, and individualised support. This is evident in the use of system-guided stages in online writing tools (Chew et al., 2020), inquiry prompts in online reading environments (Erni et al., 2023), and asynchronous grammar practice in blended secondary classrooms (Shukor et al., 2025). From an ESL standpoint, this reflects a pragmatic orientation: digital scaffolding is adopted not for its novelty, but for its capacity to extend instructional reach under structural constraints. However, this pragmatism also introduces a tension. When scaffolding is overly embedded in fixed systems or linear task sequences, it risks becoming procedural rather than contingent thus limiting responsiveness to individual learner trajectories.

Importantly, synthesising RQ1–RQ3 surfaces a persistent disconnect between sociocultural theory and classroom enactment. While scaffolding is theoretically grounded in interaction, adaptivity, and gradual fading of support, several studies operationalise scaffolding as static digital aids or pre-designed sequences that prioritise manageability and scalability. This is particularly evident in examination-driven ESL contexts, where instructional efficiency often takes precedence over responsiveness. Peer-mediated scaffolding (Tinggie et al., 2023) stands out precisely because it reintroduces interactional contingency and shared agency, aligning more closely with sociocultural principles. Yet even here, its effectiveness is contingent on learner proficiency balance, classroom norms, and teacher mediation.

These factors vary considerably across ESL settings. Conclusively, this suggests that digital scaffolding alone cannot guarantee adaptive support, instead its pedagogical value depends on how technology, task design, and social interaction are aligned within specific instructional ecologies.

CONCLUSION

This systematic review set out to examine how digital scaffolding is conceptualised, justified, and implemented within ESL contexts. Synthesising findings across the three research questions, the review demonstrates that digital scaffolding is not framed primarily as skill-specific intervention, but as a context-responsive pedagogical strategy designed to mediate structural and instructional constraints common in ESL environments. Across both secondary and higher education settings, scaffolding is mobilised to bridge misalignments between task demands and learner readiness, particularly in contexts characterised by mixed proficiency levels, limited instructional time, and examination-driven curricula.

The review further shows that digital scaffolding is predominantly realised through design-mediated and system-supported mechanisms, such as staged task sequences, guided prompts, and asynchronous practice opportunities. While these approaches extend instructional reach and stabilise learning processes, they also risk reducing scaffolding to procedural support when adaptivity, interaction, and fading are insufficiently foregrounded. Peer-mediated and interactive forms of scaffolding, though less frequently implemented, emerge as more closely aligned with sociocultural principles, highlighting the importance of agency and contingency in effective scaffolded learning.

In conclusion, the findings suggest that the pedagogical value of digital scaffolding in ESL contexts lies not in the adoption of specific technologies, but in the intentional orchestration of supports that are sensitive to learner development and contextual constraints. Future research should therefore move beyond tool-centred evaluations toward examining how digital scaffolding can be designed to remain adaptive, temporary, and interactionally grounded across diverse ESL settings.

STATEMENT ON THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE

During the preparation of this manuscript, the researchers used ChatGPT (OpenAI) solely for language editing and improving the clarity of writing. No generative artificial intelligence tools were used to produce or interpret any scientific content. After using the tool, the authors carefully reviewed and revised the text as necessary and take full responsibility for the final content of this publication.

REFERENCES

Ag-Ahmad, N., Mohamed, A. T. F. S., & Majilang, D. B. (2025). Voices from ESL classrooms: Overcoming challenges and enhancing English language teacher education in Malaysia. *Studies in English Language and Education*, 12(1), 328–345. <https://doi.org/10.24815/siele.v12i1.38886>

Adickalam, E. R., & Yunus, M. M. (2022). Systematic literature review: Investigating speaking challenges among ESL learners during the COVID-19 pandemic. *Theory and Practice in Language Studies*, 12(6), 1145–1156. <https://doi.org/10.17507/tpls.1206.14>

Budianto, L., Prihadi, K. D., Susanto, D., Rohman, A., Kholil, A., & Firdousi, M. A. (2025). Digital media's role in overcoming anxiety, enhancing linguistic elements and fostering motivation for developing speaking skills. *International Journal of Evaluation and Research in Education*, 14(2), 1379–1388. <https://doi.org/10.11591/ijere.v14i2.31931>

Chen, R., Ng, B. S., & Paramasivam, S. (2023). Toward the implications of technology-based education and EFL learners' anxiety for instructions. *Language Related Research*, 14(3), 265–287. <https://doi.org/10.29252/LRR.14.3.11>

Chuane, Q., Shukor, S. S., Yuehong, T., & Xiaofen, Z. (2023). The relationship between motivation and English language test performance among secondary vocational schools' students in China. *Studies in English Language and Education*, 10(1), 280–302. <https://doi.org/10.24815/siele.v10i1.25741>

Ding, D., & Muhyiddin Yusof, A. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study. *Humanities and Social Sciences Communications*, 12(1). <https://doi.org/10.1057/s41599-025-05550-z>

Hasan, M., & Bidin, S. J. (2023). Scaffolding students' descriptive writing skills in EFL context: A pedagogical approach. *Eurasian Journal of Applied Linguistics*, 9(3), 196–206. <https://doi.org/10.32601/ejal.903017>

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). *The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers*. *Education for Information*, 34(4), 285–291. <https://doi.org/10.3233/EFI-180221>

Kernagaran, V., & Abdullah, A. (2022). A systematic literature review of flipped learning in English as a second language (ESL) context. *International Journal of Instruction*, 15(2), 793–814. <https://doi.org/10.29333/iji.2022.15243a>

Li, D., & Zhao, Y. (2025). Artificial intelligence applications for oral communication skills in EFL contexts: A systematic review. *Asia-Pacific Education Researcher*. <https://doi.org/10.1007/s40299-025-01023-8>

Ma, M., Noordin, N., & Razali, A. B. (2022). Effect of educational technology on students' foreign language anxiety: A thematic literature review. *World Journal of English Language*, 12(6), 361–369. <https://doi.org/10.5430/wjel.v12n6p361>

Ng, M. L., Rahman, Md. M., Rahman, M. H., Lai, T. T., & Tan, R. S.-E. (2025). Exploring obstacles and effectiveness of task-based language teaching (TBLT) approach in EFL speaking instruction: Insights from an adult EFL classroom. *World Journal of English Language*, 15(6), 151–162. <https://doi.org/10.5430/wjel.v15n6p151>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>

Ramalingam, S., Yunus, M. M., & Hashim, H. (2022). Blended learning strategies for sustainable English as a second language education: A systematic review. *Sustainability*, 14(13), Article 8051. <https://doi.org/10.3390/su14138051>

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, 13(3), 273–304. https://doi.org/10.1207/s15327809jls1303_2

Shukor, S. S., Chuane, Q., Albakri, I. S. M. A., Madzlan, N. A., & Gopal, R. (2025). STARTeCH for asynchronous interactive grammar learning: Transforming ESL mastery in Malaysian schools. *Studies in English Language and Education*, 12(1), 380–397.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Walqui, A. (2006). Scaffolding instruction for English language learners: A conceptual framework. *International Journal of Bilingual Education and Bilingualism*, 9(2), 159–180. <https://doi.org/10.1080/13670050608668639>

Wang, J., Abdullah, R., & Leong, L.-M. (2022). Studies of teaching and learning English-speaking skills: A review and bibliometric analysis. *Frontiers in Education*, 7, Article 880990. <https://doi.org/10.3389/feduc.2022.880990>

Wang, N., Ahmad, N. K., & Saharuddin, N. (2025). Effective measures to develop undergraduates' communicative competence in English as a foreign language: A systematic review. *International Journal of Learning, Teaching and Educational Research*, 24(2), 416–438. <https://doi.org/10.26803/ijlter.24.2.21>

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>

Zhangli, W., Ismail, L., Ahmad, N. K., Mahfoodh, O. H. A., & Xuewen, W. (2024). Effects of blended learning on EFL college students' oral communicative competence: A critical review of studies. *Journal of Language Teaching and Research*, 15(5), 1500–1509. <https://doi.org/10.17507/jltr.1505.11>