



Review Article

A Scoping Review of Generative AI's Dual Role as Scaffolding or Substitution in Self-Regulated Learning and Learner Autonomy in EFL Contexts

Nguyen Thi Bich Ngoc¹ , Nguyen Xuan Hong^{2,*}

¹Academic Affairs Department, Phuong Nam Language Center, Ho Chi Minh City, Vietnam

²Faculty of Foreign Languages, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

Abstract

Generative artificial intelligence (GenAI) is increasingly embedded in English as a Foreign Language (EFL) instruction, with self-regulated learning (SRL) and learner autonomy positioned as key outcomes of interest. A growing evidence base reports positive effects of GenAI on self-regulatory strategies, alongside a parallel and largely separate literature warning that GenAI use may foster dependency, cognitive offloading and the externalization of self-monitoring. This scoping review maps this apparent paradox, in which GenAI appears simultaneously as scaffolding and as substitution, across the SRL and learner-autonomy literature, with particular attention to its implications for EFL learners. Following a PRISMA-ScR-informed approach, 12 empirical and review studies published between 2025 and 2026 were thematically synthesized alongside four foundational theoretical works. As a result, four thematic clusters emerged: (1) evidence of GenAI functioning as scaffolding for SRL; (2) evidence of GenAI functioning as substitution, fostering dependency and cognitive offloading; (3) design and timing features that moderate which role GenAI plays; and (4) existing reviews and meta-analyses, together with their methodological limitations. Across clusters, a consistent pattern emerges, in that nearly all studies rely on short-term or cross-sectional designs, and several recent reviews explicitly call for research measuring learners' delayed, unsupported performance so as to distinguish durable internalization from temporary dependency. No identified study has applied a longitudinal design incorporating a structured AI-withdrawal phase to directly test this distinction among EFL learners. This review argues for such a design and outlines its theoretical and methodological implications for Vietnamese EFL contexts.

Keywords

Generative AI, Self-regulated Learning, Learner Autonomy, Scaffolding, AI Dependency, EFL

1. Introduction

Self-regulated learning (SRL) and learner autonomy are widely regarded as critical determinants of sustained success in second-language acquisition, particularly once learners move beyond direct instructor supervision [16]. As generative

AI (GenAI) tools become embedded in EFL instruction, a growing body of research reports that GenAI can strengthen learners' planning, monitoring and self-evaluation strategies when appropriately designed [1, 7, 12, 14]. At the same time,

*Correspondence: Nguyen Xuan Hong (nguyenxuanhong@iuh.edu.vn)

Received: 5 September 2026; **Accepted:** 15 September 2026; **Published:** 28 September 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

a parallel and increasingly prominent literature warns that GenAI use may foster dependency, cognitive offloading and the externalization of self-monitoring processes that learners would otherwise need to perform themselves [3, 8].

This apparent paradox, in which GenAI functions on one hand as scaffolding that supports the internalization of self-regulatory competence and on the other hand as substitution that displaces it, sits at the center of current debate, yet has rarely been mapped systematically for the EFL context specifically. The stakes of resolving it are not purely academic. Vietnamese higher education has, since 2023, seen rapid and largely unregulated adoption of GenAI writing and conversation tools among EFL undergraduates, often without accompanying institutional guidance on appropriate use. If these tools function primarily as substitution rather than scaffolding for this population, the long-run cost would be measured in weakened autonomous language-learning capacity precisely at the point when learners most need it, namely after graduation, when structured instruction ends and self-directed maintenance of English proficiency becomes the learner's own responsibility. Understanding which conditions push GenAI toward one role or the other is therefore a practical, curriculum-relevant question and not only a theoretical one.

Most of the empirical base reviewed below originates from mainland Chinese, Iranian and Korean university settings. Vietnamese EFL learners, who study English under distinct curricular structures, high-stakes standardized testing pressures (notably the national high school graduation examination and university entrance requirements) and comparatively uneven technology access between urban and rural institutions, are essentially absent from this literature. This scoping review addresses that gap, guided by three review questions. First, what evidence supports GenAI's role as scaffolding for SRL and learner autonomy? Second, what evidence supports GenAI's role as substitution, fostering dependency rather than internalized competence? Third, what methodological approaches have been used to distinguish these two roles, and which approach remains unexplored? The remainder of this review proceeds as follows. Section 2 outlines the theoretical framework. Section 3 details the review method. Section 4 presents the thematic synthesis. Section 5 discusses the synthesis and identifies the resulting research gap; and Sections 6 and 7 present the review's limitations and conclusion respectively.

2. Theoretical Framework

This review is organized around three theoretical anchors, namely the cyclical model of self-regulation [16], self-determination theory [2] and the concept of scaffolding [9], elaborated through the analysis of contingent tutoring [10]. Each anchor is treated in depth below, not simply as a label for a construct, but as a set of internal mechanisms that generates specific, testable predictions about what should be observed if GenAI is functioning as scaffolding versus as substitution.

Section 2.4 then integrates the three anchors into a composite theoretical signature that directly informs the design of the longitudinal study proposed in Section 5.

2.1. Zimmerman's Cyclical Model and the Neglected Self-Reflection Phase

This model [16] conceptualizes self-regulation as a recursive cycle of three phases rather than a fixed trait. The forethought phase precedes effort and comprises task analysis processes, such as goal setting and strategic planning, together with self-motivational beliefs, including self-efficacy, outcome expectations and task interest. The performance phase occurs during the learning task itself and comprises self-control processes, such as self-instruction and attention focusing, together with self-observation processes, chiefly metacognitive monitoring and self-recording of progress. The self-reflection phase follows task completion and comprises self-judgment, meaning self-evaluation against a standard and causal attribution of the result, together with self-reaction, meaning the learner's satisfaction or dissatisfaction and the adaptive or defensive inference drawn from it. Critically, the cycle is closed rather than linear, in that the adaptive inferences generated during self-reflection become the goals and strategic choices that open the forethought phase of the next learning episode. A learner who attributes a weak draft to a specific, controllable strategy failure is positioned to revise that strategy next time, whereas a learner who attributes it to an uncontrollable cause, such as low ability, is not.

This closed-loop structure exposes a methodological weakness that runs through nearly all of the GenAI-and-SRL studies reviewed in Section 4. Because these studies measure strategy use through self-report instruments or interaction logs collected while GenAI support remains available, they capture the forethought and performance phases reasonably well, for instance through the metacognitive-support manipulations in [14] and [1], but they rarely capture the self-reflection phase under conditions where the learner must actually rely on whatever was internalized. A learner can appear, during the performance phase, to be executing a metacognitive strategy competently simply by following the GenAI tool's prompts in real time, without that execution reflecting an internalized capacity that would survive into a subsequent, unsupported self-reflection phase. Zimmerman's model therefore implies that GenAI-supported SRL cannot be adequately assessed from performance-phase data alone. The self-reflection phase, evaluated once support is unavailable, is the phase in which internalization, as distinct from momentary compliance with the tool's structure, would actually become visible.

2.2. Self-Determination Theory and the Internalization Continuum

Self-determination theory [2] grounds learner autonomy in

the satisfaction of three basic psychological needs, namely autonomy, competence and relatedness, and in the closely related organismic integration sub-theory, which describes how externally initiated behaviors become, or fail to become, part of the self over time. This sub-theory specifies a continuum of regulatory styles running from external regulation, in which behavior is performed solely to meet an external demand or secure a reward, through introjected regulation, in which the demand has been partly internalized as a source of guilt or ego-protection rather than genuine endorsement, to identified regulation, in which the learner consciously values the behavior as personally important, and finally to integrated regulation, in which the behavior coheres fully with the learner's broader values and sense of self. Only identified and integrated regulation, along with intrinsic motivation proper, count as autonomous in self-determination theory's technical sense. Introjected and external regulation remain controlled, even though both, unlike amotivation, involve some level of engagement with the activity.

This continuum offers a more precise lens on the scaffolding-versus-substitution paradox than the basic-needs framework alone. GenAI use is itself a regulated behavior, and its position on the internalization continuum, not merely its frequency, plausibly distinguishes a learner for whom the tool is scaffolding self-regulatory development from one for whom it has become a substitute for it. A learner who has identified with strategic GenAI use, in the sense of consciously endorsing it as one component of a personally chosen learning approach, is using the tool in a manner consistent with autonomous regulation and, by extension, with the kind of internalized competence that scaffolding is meant to produce. A learner whose use is driven by introjected pressure, such as anxiety about submitting an imperfect draft or guilt at the thought of writing unaided, or by external pressure, such as a course requirement, is engaging in controlled regulation that resembles the dependency construct documented in [3] far more closely than it resembles autonomous self-regulated learning. This reframing matters because the psychometric dependency instruments reviewed in Section 4.2 capture behavioral and cognitive symptoms of overreliance such as preoccupation and withdrawal discomfort, but do not directly assess the underlying motivational regulation style. Two learners with identical GenAI usage frequency could occupy very different positions on the internalization continuum, and self-determination theory predicts that only one of them is on a trajectory consistent with durable, transferable self-regulatory development.

2.3. Vygotskian Scaffolding, the Zone of Proximal Development and Contingent Fading

The zone of proximal development [9] is the distance between what a learner can accomplish independently and what the same learner can accomplish with the assistance of a

more capable other. [10], building directly on this concept, introduced the term scaffolding to describe the process by which that assistance is provided, and this analysis specifies a feature of scaffolding that is easy to overlook when the term is used loosely, namely contingency. A competent human tutor does not supply a fixed amount of help. The tutor continuously diagnoses the learner's current level of understanding from moment to moment and calibrates support accordingly, offering more direct assistance when the learner struggles and deliberately withdrawing assistance, that is, fading it, as soon as the learner demonstrates independent capability, ready to reinstate it if the learner falters again. Fading is therefore not an afterthought added once a scaffolding episode is finished. It is the mechanism that converts assisted performance into independent competence and it depends on ongoing, learner-specific diagnosis rather than on a predetermined schedule.

Read against this stricter, contingency-based definition, most GenAI tools reviewed in Section 4 depart from scaffolding in a specific and consequential way. A chatbot or writing assistant typically offers the same level of support on the tenth use as on the first, regardless of whether the learner's demonstrated competence has grown in the interim, because the tool has no mechanism for diagnosing that growth and fading its own assistance accordingly. This is precisely the structural gap identified in Section 2 of the theoretical discussion above and echoed in the design-focused findings of Section 4.3, where [4] found that a single, fixed manipulation of support timing, delayed rather than immediate, improved outcomes, and [11] found that most existing GenAI-SRL designs address only isolated phases of the self-regulation cycle rather than a genuinely adaptive, full-cycle contingent loop. Both findings are best understood as partial, externally imposed substitutes for a contingency that the GenAI tools themselves do not provide. A fixed delay rule is a coarse, non-adaptive approximation of the moment-to-moment diagnosis that defines true Vygotskian tutoring. Under this reading, whether a given instance of GenAI use functions as scaffolding or substitution depends less on the tool's inherent properties than on whether the surrounding instructional design, the course, the task sequence or the researcher's intervention, supplies the contingent fading that the tool itself structurally lacks.

2.4. A composite Theoretical Signature for the Proposed Longitudinal Design

Taken together, the three anchors generate a specific empirical question that motivates the present review and directly shapes the longitudinal, AI-withdrawal design proposed in Section 5. If scaffolding is defined by its eventual, ideally contingent removal, then GenAI's effect on SRL cannot be fully evaluated while the support remains in place. It must instead be assessed once support is withdrawn, since only performance and behavior under withdrawal can reveal whether

self-regulatory competence has been internalized or merely borrowed for as long as the tool was available.

The three frameworks jointly specify what internalization, as opposed to borrowing, should look like across three converging indicators once support is withdrawn. A scaffolding-consistent pattern would combine adaptive self-reflection in Zimmerman’s sense, meaning accurate self-judgment and constructive, controllable attributions rather than helplessness, with identified or integrated regulation of the prior GenAI use in Deci and Ryan’s sense, meaning the learner had genuinely endorsed the tool as part of their own strategy rather than leaned on it under external or introjected pressure, together with performance and strategy use that persist, or decline only mildly and temporarily, once the tool is removed, consistent with the account in [10] of competence that has been faded into independence rather than merely propped up. A substitution-consistent pattern would instead combine defensive or externally attributed self-reflection, controlled regulation of the prior GenAI use and a sharp collapse in performance and strategy use upon withdrawal. No single indicator reviewed in Section 4, whether a dependency score, a self-report strategy-use questionnaire or an immediate post-support performance measure, can establish this pattern on its own. Each captures one theoretical strand but not their conjunction. The dissertation project motivated by this review is therefore designed to collect all three strands from the same learners across a structured withdrawal phase, combining a Zimmerman-based coding scheme for post-withdrawal reflective writing or think-aloud protocols, a self-determination-theory-based measure of the learner’s regulatory style toward their prior GenAI use, and a direct, unsupported performance and strategy-use measure taken after withdrawal, so that the three theoretical strands can be examined together rather than inferred separately. A further, more fine-grained design question that follows directly from Section 2.3 is whether the withdrawal itself should occur on a fixed schedule, which is simpler to implement and more comparable to existing class-randomized designs such as [1], or on a performance-contingent basis, which is more faithful to the original definition of fading in [10] but requires continuous competence tracking. The proposed design favors a fixed withdrawal point for the primary group comparison, supplemented by continuous within-subject competence tracking as a secondary test of the contingent-fading prediction.

3. Method

3.1. Review Design

This review follows a scoping review approach, informed by the PRISMA extension for Scoping Reviews (PRISMA-ScR), appropriate for mapping a heterogeneous, rapidly evolving and conceptually contested evidence base rather than for producing pooled effect-size estimates as a full systematic review or meta-analysis would.

3.2. Search Strategy and Eligibility Criteria

Sources were identified through iterative keyword searches (e.g., “generative AI self-regulated learning,” “AI dependency learning,” “AI scaffolding fading,” “cognitive offloading generative AI,” “EFL learner autonomy AI”) covering publications from 2024 to 2026. Eligible sources were peer-reviewed journal articles and systematic reviews or meta-analyses, in English, addressing GenAI’s relationship to self-regulated learning, learner autonomy, dependency or cognitive offloading in educational contexts. Sources addressing AI in education without a self-regulation, autonomy or dependency component were excluded, as were sources published only as preprints or conference proceedings, in order to keep the evidence base restricted to peer-reviewed work.

3.3. Screening and Data Extraction

Titles and abstracts were screened for topical relevance; full texts or detailed abstracts were then reviewed to confirm eligibility and extract study design, sample, theoretical framework and reported outcomes. Twelve empirical or review studies, published between 2025 and 2026, met the inclusion criteria and form the basis of the thematic synthesis in Section 4. Four additional foundational theoretical works [2, 9, 10, 16], pre-dating the eligibility window, were retained to ground the theoretical framework in Section 2.

4. Results

The 12 eligible studies were grouped into four thematic clusters, namely evidence of GenAI as scaffolding (Section 4.1), evidence of GenAI as substitution (Section 4.2), design and timing features that moderate this balance (Section 4.3) and existing reviews and meta-analyses together with their limitations (Section 4.4). Table 1 summarizes the 12 studies by cluster, design, sample and theoretical anchor.

Table 1. Summary of the 12 included studies.

Study	Design	Sample	Cluster
[14]	Quasi-experiment	68 college students	4.1 Scaffolding

Study	Design	Sample	Cluster
[7]	Quasi-experiment	61 EFL students	4.1 Scaffolding
[1]	Class-randomized study	310 undergraduates	4.1 Scaffolding
[5]	Design and implementation study	University students	4.1 Scaffolding
[12]	Survey, structural equation modeling	386 Chinese EFL learners	4.1 Scaffolding
[3]	Scale development, six studies	1,333 participants (US and Singapore)	4.2 Substitution
[8]	Systematic review	21 studies (music education)	4.2 Substitution
[4]	2 × 2 experiment	107 students (physics tasks)	4.3 Design and timing
[15]	Survey, structural equation modeling	736 Chinese university students	4.3 Design and timing
[6]	Systematic review	18 studies (2009 to 2024)	4.4 Existing reviews
[13]	Meta-analysis	35 empirical studies (2013 to 2025)	4.4 Existing reviews
[11]	Systematic review	73 articles	4.4 Existing reviews

4.1. GenAI as Scaffolding for SRL Development

Several controlled and quasi-experimental studies report that GenAI, when paired with explicit metacognitive support, strengthens learners' self-regulatory strategies. A quasi-experiment with 68 college students found that a metacognitive support framework embedded in GenAI use significantly enhanced task strategy and self-evaluation, although without a corresponding difference in achievement, [14] a dissociation that itself foreshadows the central question this review takes up, namely whether GenAI is functioning as scaffolding or as substitution. A similar study found that GenAI-empowered interactive support strengthened university EFL students' self-regulated reading strategy use and engagement, using a purpose-built chatbot, Reade, in a quasi-experimental comparison against a one-way, non-interactive support condition [7]. The bidirectional, conversational format of the support, rather than its mere presence, appeared central to the strategy gains observed.

A class-randomized study of 310 undergraduates found that a feedback-oriented GenAI workflow (draft-first, AI critique, human evaluation) improved cognitive, metacognitive, and resource-management strategies alongside language achievement [1]. However, the authors explicitly flagged the risk of cognitive offloading absent clear usage guardrails, a caveat that anticipates the substitution literature reviewed in Section 4.2. Further evidence comes from a GenAI-based personalized recommender system, SRLAdvisor, improved both SRL and academic performance by delivering timely, individualized feedback that earlier, more labor-intensive SRL interventions such as self-reflective reports could not feasibly provide at scale [5].

Most directly relevant to the present review's EFL focus, [12] tested a chain-mediation model with 386 Chinese university EFL learners and found that the perceived ease of use, usefulness and interactivity of GenAI predicted both learners'

intention to use the technology and their subsequent learning engagement, which in turn predicted self-regulated learning. This finding is important for two reasons. It confirms that the scaffolding effect documented in non-EFL samples generalizes, at least cross-sectionally, to EFL writing contexts, and it identifies engagement, rather than mere exposure or intention, as the proximal mechanism through which GenAI use translates into SRL gains, a mechanism that has direct implications for how a future withdrawal-based design (Section 5) might be timed relative to peak engagement.

Across these five studies, a consistent pattern is that the scaffolding effect is never automatic; it appears only when GenAI use is deliberately structured around a metacognitive prompt, a feedback loop or a personalization mechanism, rather than left as unstructured, open-ended access to a chatbot. Two of these studies both built explicit metacognitive scaffolds into otherwise similar GenAI workflows, and both reported strategy gains [1, 14]. Neither study, however, reported a corresponding gain in unsupported performance once the scaffold was removed, because neither design included a withdrawal phase. This absence recurs across every study in this cluster and is taken up directly in Section 5.

4.2. GenAI as Substitution and the Risk of Cognitive Offloading

A distinct but overlapping literature documents the risk that GenAI displaces, rather than develops, self-regulatory capacity. [3] developed and validated a multidimensional Generative AI Dependency Scale across six studies, establishing dependency as encompassing cognitive preoccupation, negative consequences and withdrawal symptoms, and thereby providing a psychometrically robust instrument directly relevant to measuring the substitution side of the paradox. Notably, the scale was also shown to correlate with reduced self-concept

clarity and increased procrastination, outcomes that align conceptually with an erosion of self-regulatory capacity rather than a neutral technology-use pattern.

A systematic review of 21 studies in music education concluded that whether AI supports learner agency or erodes it depends on whether AI feedback is embedded within an internalizable learning loop or remains at a surface level of answer provision, with the latter increasing the risk that learners come to treat the system, rather than their own judgment, as the final authority on quality [8]. This finding directly parallels concerns raised for language learning contexts regarding the externalization of self-monitoring, and it suggests that the scaffolding-versus-substitution distinction may hinge less on whether GenAI is used at all than on the specific loop structure through which its feedback is delivered, a design-level variable examined further in Section 4.3.

Read against Section 4.1, this cluster suggests that scaffolding and substitution are not two different populations of tools but two different uses of structurally similar ones. The same conversational, on-demand quality that [7] and [12] associate with stronger SRL strategy use is, in the framework of [3], precisely what generates cognitive preoccupation and withdrawal symptoms when it becomes habitual rather than task-bounded. Neither cluster of studies, taken alone, can establish where the boundary between beneficial use and dependency actually lies for a given learner; only a design that manipulates the withdrawal of support directly, as proposed in Section 5, can locate that boundary empirically rather than inferring it from correlational dependency scores.

4.3. Design and Timing Features that Moderate the Scaffolding-Substitution Balance

A smaller set of studies examines which specific design choices push GenAI use toward scaffolding rather than substitution. Findings converge on the timing and structure of support as a key moderating factor. [4], in a 2×2 experiment with 107 students completing physics problem-solving tasks, crossed support timing (immediate versus delayed) with chatbot response design (open versus scaffolded) and found that delayed support produced better learning outcomes than immediate support, while response design had no significant independent effect; at the process level, immediate support was linked to learners jumping straight from prompting to adopting the AI's answer, whereas delayed support encouraged a fuller attempt at the problem first. This suggests that when support is delivered matters at least as much as how it is structured. [15], drawing on a sample of 736 Chinese university students, similarly found that AI-assisted environments shape self-regulation, autonomy and self-directed behavior differently depending on how adaptively support is calibrated to the learner's developing competence, with structural equation modeling indicating that engagement, rather than raw exposure to AI, mediated these effects.

The systematic review of 73 articles on GenAI-supported

SRL design in [11] offers a complementary, mechanism-level account, identifying six pedagogical affordances (including personalized goal setting and resource searching and integration) through which GenAI is typically embedded in SRL support, and noting that most existing designs address only isolated SRL phases rather than the full forethought-performance-reflection cycle. Taken together, these design-focused studies suggest that timing, feedback-loop structure and phase coverage are the principal levers through which GenAI can be engineered toward scaffolding, yet none of them tests whether a design incorporating explicit, staged withdrawal of support, the fading mechanism central to the original formulation in [9], changes learners' unsupported performance once the tool is removed.

4.4. Existing Reviews and Meta-Analyses, and Their Limitations

At least four prior reviews or meta-analyses bear directly on this topic. [6] conducted a systematic review of 18 studies (2009 to 2024) on AI tools and learner independence and self-regulation in language education, concluding that AI can support autonomy when integrated thoughtfully but noting the predominance of short-term study designs. [8], discussed above, reached a parallel conclusion in music education. [11], also discussed above, reviewed 73 GenAI-and-SRL design studies and likewise found the evidence base concentrated on immediate, in-session outcomes rather than on what learners can do once the support is removed, a pattern that recurs in the meta-analysis discussed next.

Most directly relevant to the present review's argument, the meta-analysis of 35 empirical studies (2013 to 2025) in [13], applying the cyclical SRL model [16] to AI-supported learning, explicitly recommends that future research examine learners' delayed, unsupported performance alongside their immediate, supported gains, in order to confirm that AI fosters genuine self-regulatory competence rather than a temporary, performance-only enhancement. This recommendation effectively calls for the withdrawal-based design this review argues for, without any included study yet implementing it. None of these four reviews focuses specifically on EFL learners, and none reports a study incorporating a structured AI-withdrawal phase. Their convergence, arrived at independently across language education, music education and general SRL design research, considerably strengthens the case that the methodological gap identified in Section 5 is not an artifact of this review's particular search strategy but a genuine, cross-domain blind spot in the current evidence base.

5. Discussion and Synthesis of the Identified Gap

Taken together, the reviewed literature supports three conclusions. First, GenAI can strengthen measurable components

of SRL when paired with explicit metacognitive scaffolding, and this effect appears to generalize to EFL writing and reading contexts specifically (Section 4.1). Second, GenAI carries a documented and psychometrically measurable risk of fostering dependency and cognitive offloading, a risk that operates through the same feedback mechanisms that, differently structured, produce the scaffolding effect (Section 4.2). Third, and most importantly, four independent reviews and meta-analyses converge on the same methodological gap. Evidence to date derives almost entirely from immediate or short-term outcomes, leaving unresolved whether GenAI-supported SRL gains persist once AI support is withdrawn, the very test that would distinguish genuine internalization, or scaffolding, from performance-only dependency, or substitution.

This gap is especially pronounced for EFL learners, where autonomy and self-regulation carry particular importance for sustained language development beyond the classroom, and where empirical evidence from Vietnamese contexts remains scarce relative to the China-dominated literature reviewed here. Of the four EFL-specific or EFL-inclusive studies identified in this review [1, 7, 12, 14], all were conducted with Chinese or Iranian university samples; none was conducted in Vietnam, and none followed learners past the period of active AI support. A longitudinal design that randomly assigns EFL learners to conditions differing only in whether GenAI support is withdrawn after an initial training phase, and then compares groups on unsupported SRL strategy use, writing or speaking performance and self-report dependency measures such as the dependency scale in [3], would directly operationalize the distinction between scaffolding and substitution that the existing literature currently only theorizes. Embedding such a design within a Vietnamese EFL setting would additionally address the geographic concentration noted above, contributing evidence from a context in which curricular structure, assessment pressure and patterns of technology access differ meaningfully from the Chinese university settings that dominate the current evidence base.

5.1. Theoretical and Practical Implications

Theoretically, confirming a withdrawal-sensitive distinction between scaffolding and substitution would extend the fading concept in [9], developed for human tutoring relationships, to a class of tools that, as noted in Section 2, has no built-in fading mechanism of its own. This would reframe the design problem facing EFL instructors and edtech developers alike, from asking whether to permit GenAI use toward asking how to engineer artificial fading into tools that would otherwise remain equally available regardless of a learner's developing competence, for instance through usage quotas that taper across a course, prompts that shift from direct answers toward guiding questions as performance improves, or periodic unsupported checkpoints of the kind proposed above.

Practically, for Vietnamese EFL programs currently deciding how, or whether, to formally integrate GenAI writing and

speaking tools into their curricula, the present synthesis suggests that the relevant policy question is not simply access versus restriction. Programs that adopt GenAI without any structured withdrawal or checkpoint component would, on the logic developed here, be adopting a design most closely aligned with the substitution-risk cluster reviewed in Section 4.2, regardless of how pedagogically framed the tool's introduction is. Programs seeking to capture the scaffolding-side benefits documented in Section 4.1 while limiting the dependency risks documented in Section 4.2 would need to build fading directly into course design, an implication that follows from the reviewed literature even before the specific longitudinal study proposed above is conducted.

5.2. Feasible Directions for Future Research

The gap identified above is broad enough that it should not be treated as calling for a single, all-or-nothing study attempted immediately after admission. A phased, practicable research program, sequenced across a standard three-to-four-year doctoral timeline, addresses it more feasibly than one long study designed to close every gap at once. Four concrete directions follow directly from the specific limitations identified in Sections 4 and 5.

Direction 1, instrument validation preceding the main study. None of the twelve studies reviewed here operationalizes the self-reflection or regulatory-style strands of the composite signature developed in Section 2.4 directly, and no ready-made instrument links the dependency construct in [3] to the identified-versus-introjected distinction developed in Section 2.2. A preliminary study, feasible within a single semester at one institution with 60 to 90 EFL undergraduates, could adapt and validate three measures before the main dissertation study is fielded, namely a Zimmerman-based coding scheme for post-withdrawal reflective writing, a self-determination-theory-based regulatory-style measure adapted specifically to GenAI use rather than to academic tasks in general and a brief unsupported performance task. Validating these instruments early keeps the larger study's measurement approach defensible rather than improvised.

Direction 2, the longitudinal structured-withdrawal design as the dissertation's central study. Building on the validated instruments from Direction 1, a class-randomized or quasi-experimental design, following the general structure used in [1] but extended past the period of active AI support, would assign EFL learners to conditions differing only in the timing or presence of a structured withdrawal phase and would track the three converging indicators identified in Section 2.4 rather than a single outcome measure. Conducting this at one or two Vietnamese universities, ideally spanning an urban and a less urban institution, would directly address the geographic and contextual gap noted above while remaining feasible within normal doctoral fieldwork constraints.

Direction 3, a factorial comparison of fixed-schedule versus performance-contingent withdrawal. Nested within the same

data collection as Direction 2, at no added recruitment cost, a subset of learners could be assigned to performance-contingent withdrawal, meaning support is removed once a defined competence threshold is met rather than on a fixed calendar date. This would allow a direct test of the contingent-fading prediction developed in Section 2.3 against the simpler fixed-schedule comparison, turning a design choice that would otherwise be made by convention into an empirical question the dissertation itself can answer.

Direction 4, extension across EFL skill domains and learner populations. The studies reviewed in Section 4.1 concentrate on writing and reading. Speaking-focused GenAI scaffolding, relevant given the oral-proficiency emphasis of many Vietnamese English curricula and assessment frameworks, along with learners outside the university level, such as upper-secondary students preparing for the national high school graduation examination, remain unexamined by the reviewed evidence base. Once the core withdrawal-design logic has been established through Directions 1 and 2, extending it to a speaking-specific task, using paired speaking assessments taken before and after AI withdrawal, offers a feasible follow-on study rather than a separate line of research, since it would reuse the instruments and theoretical signature already developed and validated.

These four directions are deliberately sequenced rather than parallel. The dissertation proper, Direction 2, rests on instruments validated in Direction 1 and can absorb the contingent-fading test of Direction 3 without additional participants while Direction 4 is left as a natural extension for the post-degree research agenda rather than a commitment built into the dissertation timeline itself, keeping the core proposal feasible within the constraints of a doctoral program.

6. Limitations of this Review

This review was conducted through iterative keyword-based web and database-indexed searches rather than a full, deduplicated export from a single bibliographic database (e.g., Scopus, Web of Science or ERIC) with independent dual-reviewer screening, as would be expected for a full systematic review submitted to a peer-reviewed journal. Accordingly, the search cannot be considered exhaustive, and some relevant studies may have been missed, particularly non-English-language studies and studies indexed only in regional databases. Before submission for publication, the author should replicate the search directly in Scopus, Web of Science or ERIC with fully specified Boolean strings. Document a PRISMA-ScR flow diagram with exact record counts at each stage; and independently verify every citation and DOI against the original source, a step made especially important by the author-attribution and citation-completeness errors identified and corrected during the preparation of this draft. The present draft is best regarded as a strong foundation and thematic map, to be formally validated through institutional database access.

7. Conclusion

This scoping review has mapped a fast-growing but conceptually divided literature on the role of generative AI in self-regulated learning and learner autonomy. Twelve empirical and review studies published between 2025 and 2026, synthesized alongside four foundational theoretical works, show that GenAI can function simultaneously as scaffolding that strengthens planning, monitoring and self-evaluation and as substitution that fosters dependency, cognitive offloading and the erosion of self-monitoring. Four independent reviews and meta-analyses, spanning language education, music education and general self-regulated-learning design research, converge on the same limitation: nearly all of the available evidence derives from immediate or short-term outcomes measured while GenAI support remains available, leaving unresolved whether any observed gains persist once that support is withdrawn. This review has argued that the distinction between scaffolding and substitution cannot be settled from performance recorded under active support alone, and that a longitudinal design incorporating a structured AI-withdrawal phase is the most direct way to test it. Grounding that design in the cyclical model of self-regulation, self-determination theory and the Vygotskian concept of contingent fading, as developed in Section 2, supplies three converging indicators, namely adaptive self-reflection, autonomous regulation of prior GenAI use and durable unsupported performance, that together would distinguish genuine internalization from temporary, performance-only dependency. Applying such a design specifically to EFL learners in an underrepresented context such as Vietnam would address a documented geographic gap in the current evidence base, since the studies reviewed here derive almost entirely from Chinese and Iranian university settings, while also speaking directly to a curricular decision that Vietnamese higher education institutions are already facing, namely whether and how to formally integrate GenAI writing and speaking tools into EFL instruction. The phased research program outlined in Section 5.2, beginning with instrument validation and culminating in the proposed longitudinal withdrawal study, offers a feasible and theoretically grounded path toward answering that question within the scope of a doctoral dissertation.

Abbreviations

AI	Artificial Intelligent
GenAI	Generative Artificial Intelligence
EFL	English as a Foreign Language
SRL	Self-Regulated Learning

Author Contributions

Nguyen Thi Bich Ngoc: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology

Nguyen Xuan Hong: Resources, Supervision, Validation, Writing – original draft, Writing – review & editing

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Chen, J., & Alibakhshi, G. (2026). AI-powered applications' effects on English language learners' cognitive, metacognitive, and resource management strategies, and language achievement. *Journal of Computer Assisted Learning*, 42(1), Article e70171. <https://doi.org/10.1002/jcal.70171>
- [2] Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
- [3] Goh, A. Y. H., Hartanto, A., & Majeed, N. M. (2025). Generative artificial intelligence dependency: Scale development, validation, and its motivational, behavioral, and psychological correlates. *Computers in Human Behavior Reports*, 20, Article 100845. <https://doi.org/10.1016/j.chbr.2025.100845>
- [4] Jiang, Y., Liu, Y., Liu, B., Feng, X.-M., & Bao, L. (2026). How does generative AI affect self-regulated learning in learners with varying knowledge levels? The role of chatbot design and support timing. *Interactive Learning Environments*. Advance online publication. <https://doi.org/10.1080/10494820.2026.2691907>
- [5] Luo, X., Wang, S., & Hew, K. F. (2026). Development and implementation of a generative AI-based personalized recommender system to improve students' self-regulated learning and academic performance. *Computers & Education*, 241, Article 105486. <https://doi.org/10.1016/j.compedu.2025.105486>
- [6] Mohebbi, A. (2025). Enabling learner independence and self-regulation in language education using AI tools: A systematic review. *Cogent Education*, 12(1), Article 2433814. <https://doi.org/10.1080/2331186X.2024.2433814>
- [7] Pan, M., Lai, C., & Guo, K. (2025). Effects of GenAI-empowered interactive support on university EFL students' self-regulated strategy use and engagement in reading. *The Internet and Higher Education*, 65, Article 100991. <https://doi.org/10.1016/j.iheduc.2024.100991>
- [8] Peng, X., Sun, K., Shan, X., & Zhang, J. (2026). Empowerment or dependency? A systematic review of the impacts of intelligent assessment and generative AI on learners' self-beliefs and cognitive agency in music education. *Frontiers in Psychology*, 17, Article 1776445. <https://doi.org/10.3389/fpsyg.2026.1776445>
- [9] Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- [10] 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>
- [11] Xia, Q., Liu, Q., Thili, A., & Chiu, T. K. F. (2026). A systematic literature review on designing self-regulated learning using generative artificial intelligence and its future research directions. *Computers & Education*, 240(1), Article 105465. <https://doi.org/10.1016/j.compedu.2025.105465>
- [12] Xu, J. (2026). How generative AI enhances self-regulated learning in EFL learners: A chain mediation model of “intention to use” and “learning engagement.” *Frontiers in Psychology*, 17, Article 1808183. <https://doi.org/10.3389/fpsyg.2026.1808183>
- [13] Xu, J., Luo, Y., Wang, C., Wang, M., & Wu, Y. (2026). AI support in self-regulated learning: A decade of technological evolution and meta-analysis. *British Journal of Educational Technology*, 57(5), 1191–1220. <https://doi.org/10.1111/bjet.70058>
- [14] Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. *British Journal of Educational Technology*, 56(5), 1842–1863. <https://doi.org/10.1111/bjet.13599>
- [15] Yang, L. (2026). Empowering the autonomous learner: How AI-assisted language learning environments shape self-regulation, autonomy, and self-directed behaviors. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688261422129>
- [16] Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.