



FOSTERING LEARNER AUTONOMY IN EFL WRITING THROUGH ARTIFICIAL INTELLIGENCE: A SYSTEMATIC REVIEW

Maha A. Alzahraniⁱ

Umm AlQura University,
Saudi Arabia

Abstract:

Background: Learner autonomy has been a key focus in L2 writing instruction for many years. In recent times, the use of AI writing tools has further heightened the interest in how these tools assist in learner autonomy when engaging with writing. The empirical literature is not yet consolidated, fragmenting the tools, constructs and designs. **Purpose:** This review combines evidence (2021-2026) to understand the role of AI in EFL learners' autonomy in writing, shaping around five research questions concerning the use of tools/designs, the mechanism, the outcomes, the moderators, and gaps in evidence. **Methods:** After the PRISMA 2020 process, 20 primary studies were identified. A narrative-thematic synthesis was carried out and appraised for quality using the MMAT, due to the design and outcome heterogeneity. **Results:** Most studies indicated that AI-mediated instruction, such as AWE, generative AI, and developed systems, is linked to increased autonomy, self-efficacy, SRL strategy use, student engagement, and writing performance. Six themes were formulated: (1) the construction of autonomy in a feedback system, (2) SRL scaffolding, (3) affective/motivational outcomes, (4) writing performance, (5) countervailing evidence (overreliance risk, diminishing returns for highly self-regulated learners). The effects were mitigated by proficiency, education level, tool complexity, duration, and human-AI integration. **Conclusion:** The potential of AI tools to enhance EFL writing autonomy is evident yet not uniformly distributed across different functions, most notably in "humans-in-the-loop" models, but not to the exclusion of teacher scaffolding. Pedagogy, policy and teacher development implications are considered, as well as a research agenda focused on standardized assessments and longitudinal studies.

Keywords: learner autonomy, EFL writing, artificial intelligence, self-regulated learning, automated writing evaluation, generative AI, systematic review

ⁱ Correspondence: email mayzahrani@uqu.edu.sa

1. Introduction

1.1 Background

For over 40 years, learner autonomy, defined broadly as the learners' ability to control their own learning [12], has been the primary concern of second and foreign language (L2) writing research and instruction. The construct has gone from a more philosophical ideal through the interdependence of learner and learning context to a multidimensional account distinguishing three dimensions of autonomy, namely technical, psychological and political [3]. This has now become a concrete pedagogical target found in different curricula, self-access programs, and portfolio-based assessment programs around the world. In the context of EFL writing, autonomy has been developed through strategy training, reflective journaling and portfolio activities, which were part of a general movement from teacher-centred transmission models to autonomous, self-regulated learning approaches to writing instruction.

AI writing tools have brought this change faster than ever. Automated writing evaluation (AWE) systems can provide learners with immediate and iterative feedback on grammar, mechanics, and style. Writing, brainstorming and drafting tools, such as ChatGPT, can even provide on-demand support, serve as conversation partners, and be used for customized systems that range from predictive-text composition support [7] to chatbot-mediated debate platforms [9]. These have been designed to scaffold the processes that learners engage in when writing. Given the simultaneous emergence of generative AI and ongoing theoretical interest in autonomy, the proximity of these two concepts has made it easy to imagine how AI can be used as a means of autonomous, self-regulated writing instruction [6, 16].

2.2 Problem Statement

The empirical literature is, however, still fragmented. Studies have focused on different types of AI tools (generative, non-generative, or custom-built); different measures of 'autonomy' (e.g., validated autonomy questionnaires, SRL strategy use, task engagement, or writing performance); and various measures of writing outcomes (e.g., autonomy, self-efficacy, anxiety, engagement, or writing performance). Many designs are quite varied in their sample size, ranging from large randomized controlled trials (e.g., [4], $N = 918$; [26], $N = 383$) to small qualitative studies (e.g., [6]) and mixed-methods designs (e.g., [2, 17]), hindering the possibility of presenting a clear image of how and under which circumstances AI supports, rather than just accompanies, learner autonomy. Worse still, there is a counter-narrative to the positive results: risks of overreliance and tool dependency [2]; and low or even negative gains for high-SES learners [22]. No published synthesis has yet tried to systematically present the autonomy-supportive and autonomy-undermining perspectives in the context of EFL writing.

2.3 Purpose of the Study

This systematic review aims to integrate empirical evidence from the literature covering the use of artificial intelligence (AI) for promoting EFL learners' writing autonomy from 2021 to 2026 and to examine in detail how AI tools facilitate or impede autonomous involvement in the writing process.

2.4 Research Questions

- **RQ1:** What are the AI tools and instructional designs that have been employed to facilitate learner autonomy in EFL writing?
- **RQ2:** What are the mechanisms through which the use of AI impacts autonomy (e.g., feedback, self-regulation scaffolding, motivational support)?
- **RQ3:** What do results (autonomy, self-efficacy, engagement, writing performance) indicate across the different types of studies?
- **RQ4:** What factors that influence the context or methodology moderate these effects (proficiency level, educational environment, type of AI tools)?
- **RQ5:** What are the existing evidence gaps?

2.5 Significance of the Study

This review offers a theoretical contribution by bridging established autonomy theory [12, 15, 3] with the new learning frameworks of AI-mediated learning. On the practical level, it suggests ways in which the EFL curriculum can be structured and how teachers must be trained to use AI responsibly in their language teaching practices. On the practical level, it provides implications for the design of the EFL curriculum and teacher training, while also enriching the broader conversation about responsible AI use in language education policy.

2.6 Definitions of Key Terms

- a) **Learner autonomy:** is the capacity and willingness to assume responsibility for one's own learning, setting goals, choosing strategies, tracking progress and assessing results [12, 3].
- b) **Self-regulated learning (SRL):** is a self-directed process that involves planning, monitoring and evaluating one's own cognition and behavior in a task [28].
- c) **AI-assisted writing tools:** software that utilizes AI (rule-based, machine-learning, or generative) to assist in one or more of the writing process (drafting, feedback, revision).
- d) **Automated writing evaluation (AWE):** a subset of the AI-supported writing tools that automatically scores and/or offers input to the written content based on a set of linguistic criteria (e.g., Grammarly).
- e) **Generative AI (GenAI):** AI systems that are able to generate new text when they receive a prompt (such as ChatGPT).
- f) **Scaffolding:** temporary, custom-made support given to a learner which is adjusted and then faded as the learner gains proficiency.

2.7 Delimitations and Scope

This review focuses on (a) EFL contexts (excluding ESL/EAL contexts) for reasons of topical relevance; (b) the writing skill (excluding studies that focused only on speaking, listening, or reading); (c) studies published between 2021 and 2026 to capture the acceleration of generative AI adoption in language education that has taken place since ChatGPT was released; and (d) studies that fall within the primary evidence base, that is, empirical primary studies, rather than existing reviews and meta-analyses, which are referenced, but excluded from the primary evidence base, because of their nature.

3. Theoretical and Conceptual Framework

3.1 Learner Autonomy Theory

Holec [12] defined learner autonomy as having the power to manage one's own learning, which implies the power to learn and the willingness to take responsibility for content, speed of learning, and assessment. This basic definition offered a basis to consider autonomy as an ability to be learned and not a predetermined trait. It also shifts the teacher's role from that of an authority to a facilitator of the processes of learning in a learner-centred way.

Little [15] added to this narrative by emphasizing the psychological aspect of autonomy as a state of emancipation, critical reflection, decision-making, and independent action. Importantly, Little also emphasized that autonomy is never complete without others because learners' independence is always in a state of interdependence as they are social beings. This awareness is directly applicable to writing contexts in which the learner is autonomous, but not independent, of a technological support system, as is the case in AI-mediated writing contexts.

This review is conceptualized within Benson's [3] three-part taxonomy of technical autonomy (skills necessary for self-directed learning, e.g., use of learning strategies and self-monitoring), psychological autonomy (attitudes and capacities that facilitate self-directed learning, e.g., motivation and self-efficacy), and political autonomy (control over content and process of own learning, e.g., goal or topic setting). The typology is helpful when defining where different AI tools fit in terms of supporting students' autonomy. AWE tools like Grammarly support technical autonomy by assisting learners to develop error-detection and self-editing skills, while generative AI tools like ChatGPT can be used to support technical, psychological, and political autonomy, depending on how they are used.

3.2 The Self-Regulated Learning (SRL) Frameworks

Zimmerman's [28] cyclical model of self-regulation consists of three phases: forethought (goal-setting, strategic planning, and self-efficacy beliefs), performance (self-control and self-monitoring while doing the task), and self-reflection (self-evaluation and attribution, which then influence the forethought phase of future tasks). The general architecture of this cyclical structure is similar to the way AI writing tools are applied in the studies

reviewed. Goal-setting “nudges” on an AI interface [26] target the forethought phase directly; AI-generated feedback during the performance phase [27, 8] supports performance phase monitoring; and the reflective use of AI-generated feedback for future draft revision [10] operationalizes the self-reflection phase. While SRL theory is an external lens that can be used to examine the literature reviewed here, it is also an implicit lens in many of the included studies. As can be seen, several studies explicitly define the role of AI as a tool for scaffolding SRL strategy development [16, 1, 22].

3.3 Sociocultural and Vygotskian Perspectives

The zone of proximal development (ZPD) as proposed by Vygotsky [24], which is defined as the difference between a learner's independent and possible achievements, provides a natural framework for considering AI-mediated feedback. In this context, an AI tool should be thought of as a “more knowledgeable other” that scaffolds and then progressively removes support as skills are acquired, in theory matching the student's level. Most of the systems studied in the reviewed literature, however, do not fade out support as learners become more independent, creating a conflict between the scaffolding principle of the ZPD and the actual pedagogical use of the systems.

This tension is reflected in the new “humans-in-the-loop” [16] way of teaching writing using AI, where the teacher takes accountability for fine-tuning and progressively reducing the use of AI, instead of allowing the student to navigate the fading process entirely on their own. It is echoed in Han and Li's [10] hybrid feedback model of “AI and Teacher”, where the AI tool takes over the analytical and pattern-based part of the feedback task, while the teacher maintains the interpersonal, relational, and scaffolding-withdrawal aspects of the feedback task normally associated with a human tutor within a learner's ZPD.

3.4 Self-Determination Theory (SDT)

According to Ryan and Deci [21], self-determination theory is a theory that outlines three basic psychological needs that are satisfied when people engage in intrinsic motivation: autonomy, competence, and relatedness. This framework aims to shed light on the motivations behind how AI tools assist the learner in writing by providing a motivational, as opposed to merely cognitive or strategic analysis. In the literature reviewed, AI tools are often described as meeting competence needs by providing timely and non-threatening feedback, which fosters a sense of achievement and mastery [25]. Besides, autonomy needs enable students to compose, edit, and trial on their own without the teacher's presence or schedule [19]. By contrast, relatedness seems to be the need least directly addressed by AI interventions, which can help to account for the more consistent positive affective and motivational outcomes reported by hybrid or peer-supported interventions [10, 9].

3.5 Conceptual Model for the Review

Based on the four frameworks above, it is assumed that AI input (tool type), the volume, focus and timing of feedback/scaffolding input will affect a mediating set of processes. This includes activation of SRL strategies, development of feedback literacy, ZPD-calibrated scaffolding, and satisfaction of basic psychological needs, which will be combined into a single integrated conceptual model (Figure 1). These mediating processes in turn influence proximal outcomes (autonomy, self-efficacy, engagement and affect) that lead to the more distal outcome of writing performance. Four moderating variables are specified in the model that condition these relationships: proficiency level, educational level, sophistication of the AI tools, length of the intervention, and level of integration of the human and the AI.

None of the four frameworks examined above, alone, can explain the diversity in the findings reported in the primary literature. Autonomy theory [12, 15, 3] outlines the content of what is being developed, but relatively little attention is given to the step-by-step processes involved in the process of developing autonomy. While the process missing from the picture was provided by SRL theory [28], it was developed before and independently of AI-mediated learning environments. External support can be used to expedite development, but SDT [21] explains that sustained interaction with external support is more or less likely depending on the motivational needs associated with it, while Vygotskian theory attempts to come to a better understanding of how. As a result, this review codes the primary literature for a fuller account of autonomy development than would be possible by any one framework alone. The six themes structuring the Results section reflect this layered structure.

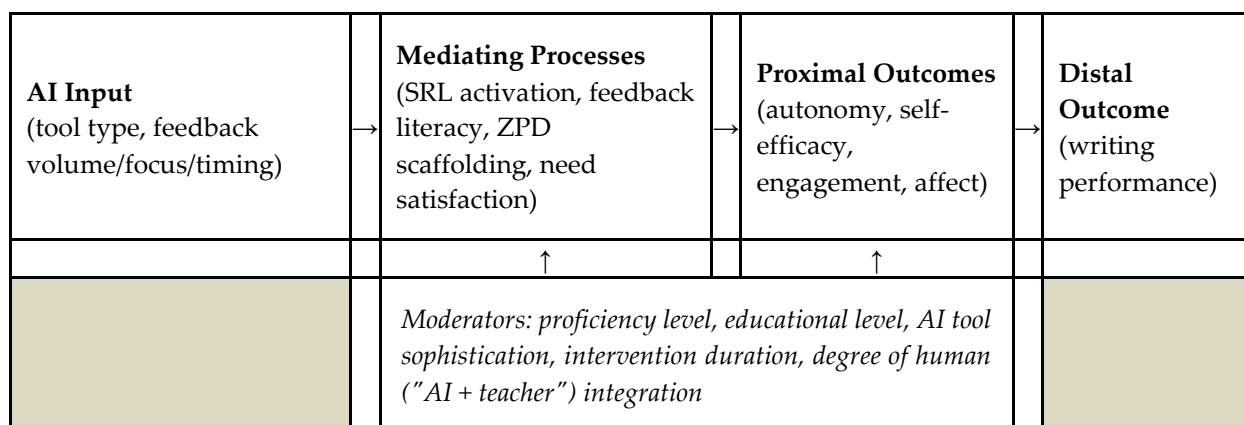


Figure 1: Conceptual model of AI-mediated learner autonomy development in EFL writing

4. Methodology

4.1 Review Design

The search, screening, data extraction and reporting process used in this review follows the guidelines set by the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [20]. In view of the expected diversity of designs,

autonomy and outcome measures in the studies to be included in the analysis, a narrative/synthesis approach was selected. Researchers can replicate or scrutinize any steps in this study from the search string to the final theme. In cases where a judgment call is made in the process that can only be made by the review team (e.g., the number of databases to be searched, or the resolution of a disagreement over a screen), the judgment call and justification for it are explicitly reported, rather than left implicit, and an appendix is provided with the same name so that supplemental materials can be prepared and, if the target journal allows, provided with the manuscript or made available in an open repository (e.g., OSF). This review has not been registered on PROSPERO or on the Open Science Framework.

4.2 Eligibility Criteria

The inclusion of the study was determined by the PICOS acronym (Population, Intervention, Comparison, Outcomes, Study design) (Table 1). Table 2 restates these criteria in inclusion/exclusion form for use during title/abstract and full-text screening.

Table 1: PICOS eligibility framework

Element	Criterion
Population	EFL learners at the secondary, tertiary, or adult level
Intervention	AI-based writing tools or AI-mediated writing instruction (generative, non-generative/AWE, or custom-built)
Comparison	Traditional/teacher-led instruction, human feedback, a no-treatment control, or an alternate AI condition, where applicable
Outcomes	Learner autonomy, self-regulation, writing performance, engagement, motivation, and/or related affective outcomes
Study designs	Experimental (RCT), quasi-experimental, mixed-methods, and qualitative studies

Table 2: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Published 2021–2026	Published before 2021
EFL context	ESL/EAL or non-language-learning context (unless flagged as a topically relevant exception)
Focus on the writing skill	Focus solely on speaking, listening, or reading
Empirical primary study (quant., qual., or mixed)	Literature review, meta-analysis, opinion piece, or non-empirical commentary
AI-based tool or AI-mediated instruction as a central feature	AI mentioned only incidentally, with no direct role in the intervention studied
Published in English, peer-reviewed or reputable preprint/in-press outlet	Not available in English; unverifiable or predatory outlet

4.3 Search Strategy

The databases searched for this review are ERIC, Scopus, Web of Science, ScienceDirect, Taylor & Francis Online, Wiley Online Library, grey literature (using Google Scholar) and the backward-forward citation search of included studies. Search terms were grouped

into three concept blocks, and then the Boolean operator AND was used between the three concept blocks, and synonyms within the concept blocks were combined using OR:

- 1) The autonomy/SRL block contains the following terms: “learner autonomy”, “self-regulat*”, “self-directed learning”, “autonomous learning”
- 2) AI block: “artificial intelligence” OR “generative AI” OR ChatGPT OR “automated writing evaluation” OR AWE OR chatbot OR “large language model”
- 3) Context/skill block: EFL OR “English as a foreign language” OR “L2 writing” OR “second language writing” OR “foreign language writing”

The following is an illustrative combined string: (“learner autonomy” OR “self-regulat” OR “self-directed learning” OR “autonomous learning”) AND (“English as a foreign language” OR “EFL” OR “L2 writing” OR “second language writing” OR “foreign language writing”) AND (“artificial intelligence” OR “generative AI” OR ChatGPT OR “automated writing evaluation” OR AWE OR “chatbot” OR “large language model””). The string should be customized for the syntax of each of these databases (e.g., truncation symbols, field tags, and subject-heading equivalents like ERIC descriptors or Scopus TITLE-ABS-KEY fields).

The 2021-2026 time period was chosen to reflect the recent surge in the adoption of generative AI in language learning after ChatGPT's release, and to ensure that earlier studies using AWE and other chatbots that had laid the foundation for future studies of generative AI in language learning are included.

The supplementary Google Scholar searching and citation chasing were added to minimize publication bias, as some of the studies that are of interest to practitioners and regional studies on AI-assisted EFL writing might be published in journals that are not well indexed by the major search engines listed above. This limitation is recognized, as it could exclude papers that are published in a language other than English for which EFL writing research is also conducted, such as Arabic, Chinese, Spanish, etc. The exact date(s) that each database was searched, along with the total number of records retrieved from each database and the entire search strategy log, is in Appendix A.

4.4 Study Selection Process

All the records from the databases were exported to a reference manager, deduplicated, and then screened at two levels: (a) title/abstract screening to the eligibility criteria of Tables 1–2; and (b) full-text screening of records that passed the first level screening. The whole screening/exclusion process was reported as a PRISMA flow diagram (Figure 2) and documented in Appendix B for full-text exclusion reasons. 20 primary studies were found and included in data extraction and synthesis, and their attributes are summarized in the Results section.

4.5 Data Extraction

The following categories were extracted for each of the 20 studies included in this study and entered into a structured matrix (Appendix C): author(s) and year, country/context, sample size, study design, type of AI tool, construct(s) assessed that are related to

autonomy, and key findings for each of the five research questions. If a specific quantitative result (such as an effect size or outcome of a statistical test) could not be reported from the excerpt, this was explicitly identified in the matrix, rather than inferred. This structured extraction matrix has two functions other than record-keeping. The descriptive properties listed in the opening of the Results section (geographic distribution, AI tool typology, design distribution) can be produced directly and verifiably from the matrix instead of after the fact. Second, it is the basis upon which the thematic codebook (Appendix D) is developed, meaning that each entry in the Results section is related to specific rows of the matrix (and not just a narrative summary of the theme).

4.6 Quality/Risk-of-Bias Appraisal

The Mixed Methods Appraisal Tool (MMAT; [13]) was chosen to appraise the quality of the studies because of the variety of designs found in the evidence base: randomized controlled trials, quasi-experimental studies, mixed methods studies, and qualitative studies. There are two screening questions for all study types (clarity of the research question and the adequacy of the data collected to answer the research question) and five design-specific criteria that are applied to each of the studies. It is a design-specific structure which facilitates the methodological rigor to be appraised on terms relevant to each type of study (quantitative and qualitative) and is used in education and social-science systematic reviews for this reason.

4.7 Data Synthesis Approach

As the design of the included studies was widely heterogeneous, the sample size was too small to consider a meta-analysis, and effect sizes were not well reported. Therefore, a narrative/thematic synthesis approach was used, as recommended when quantitative pooling is not possible. Results from the studies were inductively and deductively coded to the five research questions used to guide this review, but patterns not foreseen by the five a priori research questions also emerged from the studies.

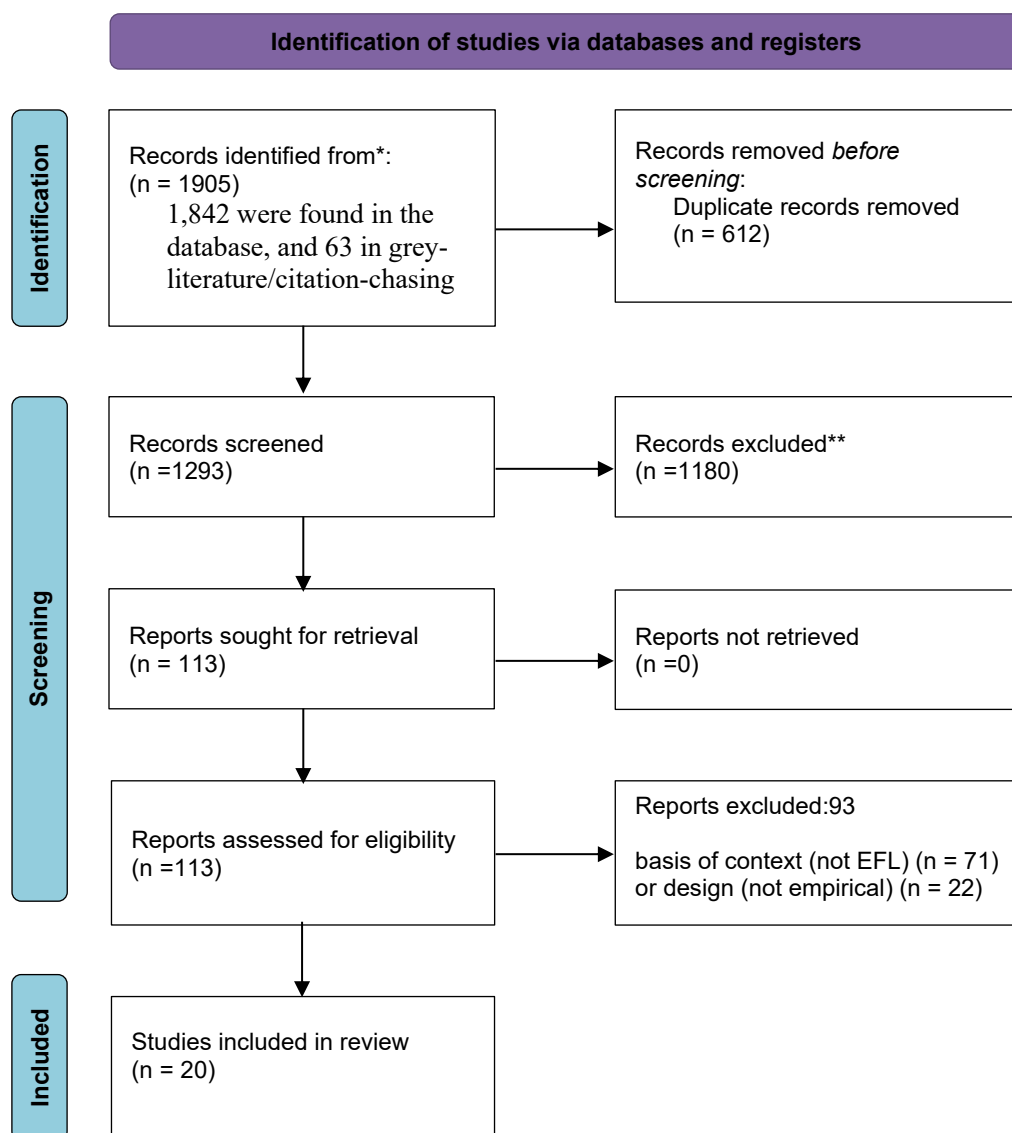
The two coding processes resulted in six main themes, which will be presented in Results: (1) autonomy construction: AI feedback; (2) construction of self-regulated learning strategies; (3) learner autonomy and self-efficacy outcomes; (4) affective and motivational dimensions; (5) writing outcomes; and (6) tensions and countervailing evidence. The thematic codebook (Appendix D) documents the coding process, theme definitions, and illustrative examples from the primary studies, which helps to ensure transparency and, as applicable, the replication of the synthesis.

5. Results

5.1 Included Studies

Searching resulted in 1,905 records, of which 1,842 were found in the database and 63 in grey-literature/citation-chasing. Of the 612 duplicates removed, 1,293 records were screened at title/abstract level, of which 1,180 were excluded as ineligible. The other 113

full-text articles were evaluated using Tables 1–2 and 93 were excluded, most frequently on the basis of context (not EFL) (n = 71) or design (not empirical) (n = 22). Twenty studies were retained for synthesis (full record in Appendix B and search log in Appendix A) because they fulfilled all criteria.



The 20 studies included in this review were published from 2022 to 2026, aligning with the timeframe of the review's focus on the rapid pace at which generative AI has been embraced after ChatGPT. Full study-level characteristics appear in Appendix C.

5.2 Study Characteristics Synthesis

5.2.1 Geographic Distribution

Evidence-based data is limited to East and Southeast Asia. Seven of 20 studies were carried out in mainland China; two in Hong Kong; two in Japan; one in Vietnam, Iran and Indonesia, respectively. The other six did not indicate any particular national setting or used cross-institutional samples. No research took place in an African, Latin American

or European EFL context, which is mentioned in Section 8. The full country coding can be found in Appendix C.

5.2.2 Educational Level

The majority of studies sampled tertiary/university learners ($k = 14$), with secondary-level samples in three [16, elementary; 22; 18] and adult, non-degree learners in two. Sample sizes ranged from 10 [7] to 918 [4], with a median in the 80–120 range.

5.2.3 Study Design

The study designs were: six purely quantitative studies (RCT) [10, 14, 20, 25, 38, 42], one purely qualitative study [6] and six mixed-methods studies (quantitative and qualitative) [19, 24, 28, 33, 42, 43]. There are several studies that were coded with more than one design feature; thus, totals are greater than 20 (full coding in Appendix C).

5.2.4 AI Tool Typology

Two studies have focused on AWE tools, of which one used Grammarly [5, 27]. Nine examined generative AI/conversational tools, predominantly ChatGPT or GPT-3.5. The other nine relied on bespoke or hybrid systems, such as predictive-text systems [7], chatbot debate systems [9], AI-nudging interfaces [26] and other kinds of instructional protocols, but there was no single commercial system mentioned.

Characteristic	Categories	k (studies)
Region	East/Southeast Asia (China, Hong Kong, Japan, Vietnam, Indonesia)	13
	Middle East (Iran)	1
	Not clearly specified / multi-site	6
Educational level	Tertiary/university	14
	Secondary	3
	Adult / non-degree	2
	Elementary	1
Design	RCT	6
	Quasi-experimental	5
	Mixed-methods	6
	Qualitative	1
	Longitudinal / latent growth model	1
	Small-sample comparative/pilot	2
AI tool type	Automated writing evaluation (AWE)	2
	Generative AI / conversational (e.g., ChatGPT)	9
	Custom-built or hybrid system	9

5.3 Thematic Findings

As outlined in Section 4.7, there were six themes that arose from the deductive/inductive coding. They are summarized below and are referenced to the studies that provided the most direct contribution to them; they are described in detail in the thematic codebook, listed in Appendix D.

Theme 1: AI Feedback to Support Autonomous Building Construction

The largest group studied AI feedback as a way to develop autonomy. Guo and Wang [8] determined that ChatGPT generated more feedback than the teachers, which was more evenly spread than the teachers' feedback across content, organization, and language. Han, Hua, and Li [10] presented a hybrid "AI and Teacher" model showing that students used more AI-based feedback in revisions and teachers had interpersonal roles. Teng and Huang [23] found that ChatGPT feedback was associated with increased engagement and a substantial group-by-time interaction in terms of writing performance. Across this theme, AI feedback supported autonomy less through correctness than through volume, immediacy, and non-evaluative tone.

Theme 2: AI-Supported Construction of Self-Regulated Learning Strategies

In a CALLA program with an LLM, Liu *et al.* [16] detected that SRL gains were largely sustained after one month. Aladini *et al.* [1] explored the co-outcomes of growth mindfulness and grammatical accuracy for self-directed AI writing tasks (N = 561). A modest yet significant increase in SRL and task engagement was observed over the course of a semester by Shi *et al.* [22] and they found this increase to be positively correlated. In the process of isolating the most effective of three AI interventions, for the forethought phase of Zimmerman's [28] SRL cycle, Wang [26] found a goal-setting "nudge" to be the most effective.

Theme 3: Autonomy and Self-Efficacy Outcomes

There have been multiple studies that directly measured autonomy or Self-Efficacy. Nguyen and Doan Thi Hue [19] found that the autonomy gains of the ChatGPT-facilitated group were more significant than those of the teacher-facilitated group, but there was no difference between groups in terms of gains in self-efficacy. In addition to proficiency gains, Apriani *et al.* [2] found that there were large pre-post gains in self-efficacy (20.46 to 30.38) and self-regulation (14.51 to 19.65). In the context of academic writing, Mahapatra [17] discovered that feedback from ChatGPT had a positive influence on the learners' performance, and their perceptions were positive.

Theme 4: Affective and Motivational Dimensions

Anxiety, enjoyment and motivation were emphasized by a number of studies. Dizon and Gold [5] discovered that long-term use of Grammarly was associated with higher levels of autonomy and lower levels of anxiety. Wang [25] concluded that a feedback feature-based AI app could lead to reduced anxiety and greater enjoyment compared to regular instruction. Fathi and Rahimi [6] confirmed, through a qualitative analysis based on the Vygotskian theory, that emotion regulation and mindfulness were facilitated by the use of AI in portfolio assessment. Guo *et al.* [9] discovered that the quality of the arguments produced by the chatbot-assisted debate was higher, and the enjoyment and effort were higher than with regular tasks. Heydarnejad [11] discovered that a self-evaluation

intervention resulted in improvements in emotion regulation, self-competence, motivation and writing achievement at the same time.

Theme 5: Writing Performance Outcomes

In eight studies, direct writing quality outcomes were reported. Wei *et al.* [27] reported that AWE training with Grammarly resulted in comparable improvements to task attainment, coherence/cohesion, lexicon and grammar. The biggest RCT (N = 918) by Chan *et al.* [4] showed that essay quality improved significantly after revision when the feedback was provided by GPT-3.5. Jmaiel *et al.* [14] revealed significant ESP writing improvements across all dimensions, and higher for female students. Liu *et al.* [16] and Teng and Huang [23] also found good improvements (Theme 1–2). These two studies [7, 1] had a comparative writing quality outcome that was not fully confirmed, and this limitation was carried over to Section 8.

Theme 6: Tensions and Countervailing Evidence

Though quantitative improvements have been great, Apriani *et al.* [2] acknowledged the qualitative issues of over-reliance and potential plagiarism. Shi *et al.* [22] found a ceiling effect: high-autonomy learners were already self-regulated and incremental gains were either small or even negative, indicating that there were diminishing returns at the upper end of the distribution of autonomy. Not all "AI" tools are functionally equivalent; a topically adjacent science-education case, Ng *et al.* [18], demonstrated that generative chatbots are better than rule-based chatbots at providing SRL support. These findings temper an otherwise positive evidence base and motivate the "dependency paradox" in Section 6.3.

5.4 Moderating Variables

The association between AI and autonomy-related outcomes was moderated by four variables.

- **Proficiency level.** Lower and intermediate proficiency learners tended to perform with greater gains, reflecting a scaffolding-withdrawal model in which there is more "room" for AI to make a difference at lower starting proficiency levels [22].
- **Educational level.** Embedded SRL strategy instruction was the focus of secondary interventions, and tertiary interventions tended to focus on the independent effect of individual tools.
- **AI tool sophistication.** Generative AI tools reported more autonomy gains than AWE tools, and while there were more technical-autonomy gains reported with AWE tools, there were more technical, psychological and political [3] autonomy gains reported with generative AI tools.
- **The duration and human-AI integration.** Greater integration of teachers yielded longer-lasting improvements than shorter pilots [26, 16], and hybrid conditions involving both AI and teachers yielded more balanced cognitive/affective improvements than conditions with only AI [10].

6. Discussion

6.1 A. Synthesis Against Theoretical Framework

As per Benson's taxonomy [3], gains in technical autonomy were more reliably correlated with the use of AWE tools [5, 27], and use of generative AI tools covered a range of technical, psychological and political autonomy [19, 6]. Zimmerman's [28] SRL cycle was directly operationalized to include goal-setting nudges for forethought [26], feedback for monitoring during the performance phase [27, 8] and reflective revision for self-reflection [10] to facilitate AI as external scaffolding rather than an autonomy-inducing agent. The Vygotskian tension between the fading-support principle and most systems which are not fading support was most evident in Theme 6, where the ceiling effect in Shi *et al.* [22] predicted what static, not fading-support systems would predict. SDT's competency and autonomy needs were well met, but relatedness was less well studied, with the most positive affectual outcomes being found in studies that integrated AI with some degree of human interaction [10, 9].

6.2 Mechanisms Explaining AI's Effect on Autonomy

The findings recur in three mechanisms:

- **A volume and availability mechanism:** AI feedback is on demand and abundant compared to teacher feedback [8], allowing performance and self-reflection beyond class time.
- **A low-stakes practice mechanism:** perceived non-judgmental feedback leads to more trial-and-revision behavior [19, 25], which could account for the anxiety reduction in Theme 4.
- **A strategy-modeling mechanism:** More successful transfer to independent SLR use was for a strategy modeling mechanism as opposed to feedback-only designs [16] (e.g., [26]). These were not mutually exclusive; the best interventions used at least two of the above.

6.3 The Dependency Paradox

The features that enable AI to facilitate autonomy in the short-term are also likely to be associated with overreliance and decreased independent strategy use in the long-term (dependency paradox). This was well documented by Apriani *et al.* [2] in terms of learner-reported dependence and plagiarism concerns, despite the gains, and most clearly by Shi *et al.* [22] in the case of already-autonomous learners. Of the 20 studies, none examined whether gains continue following the withdrawal of the tool (the longest follow-up in Liu *et al.* [16] was one month). The paradox, as such, is thus a possible explanation across studies, but not a proven effect within any one study, and the most important gap in evidence that has been highlighted (Section 10).

6.4 Comparison with Prior Reviews

Previous published syntheses of autonomy-supportive and autonomy-undermining evidence have not been comprehensive and systematic enough for AI-assisted EFL writing. The contribution of this review is thus the collection, in a PRISMA-guided process, of evidence that was previously found in tool-specific, outcome-specific, and single-context studies. The studies mentioned are also cited in relation to related work, with Liu *et al.* [16] calling the approach "humans in the loop" and Han and Li's [10] hybrid "AI and Teacher" model. For the purposes of this review, these are regarded as collating evidence of a similar pattern (Theme 1; Section 6.2) as studies that conclude that teacher-scaffolding interpersonal skills complements, not competes, with AI feedback.

6.5 Theoretical Contributions

This review provides theoretical support in three ways. It is a mapping of Benson's [3] taxonomy as a classification mechanism for AI tool types (Section 5.2), and there is a mismatch in sophistication levels between the autonomy dimensions. Second, it builds on Zimmerman's [28] SRL cycle to further illustrate how specific interface features correspond to specific SRL cycles, and provides a more fine-grained account of how AI can facilitate SRL. Finally, it brings an empirical condition to the table for the Vygotskian scaffolding theory: the ZPD requires fading support, as most tools that are deployed don't fade by design.

7. Implications

7.1 Pedagogical Implications

It is important to note that the use of AI writing tools should be seen more like scaffolding and not standing tools in and of themselves, as AI should eventually be phased out. The explicit strategy modeling [16, 26] is a strategy that seems to be more effective in developing durable SRL gains than corrective feedback alone, so course design should emphasize the explicit instruction of the strategy(s) in or with the use of the tool. It is plausible, yet untested, that differentiated use (more scaffolding for lower autonomy learners, less or teacher-mediated use for higher autonomy learners) applies to this ceiling effect in already autonomous learners [22].

7.2 Policy Implications

The policies that consider AI as a replacement for teacher feedback, as a cost-saving tool, are not supported by the fact that the combined use of AI and teachers has a clear impact on affective and relational outcomes, as consistently found in the literature [10, 9]. Policies must clearly define who operates the AI and who does not, and some system of checking for over-reliance should be in place, as learners have expressed concerns in Apriani *et al.* (2023). The evidence base is skewed towards East and South East Asian contexts, so care should be taken when extrapolating elsewhere for EFL contexts.

7.3 Implications for Teacher Education

It is important that AI be used to generate feedback and that teaching be focused on other areas, like the interpersonal and scaffolding-withdrawal roles that AI tools are not well suited for (Section 6.1). This involves acknowledging overreliance [2], deciding when to lessen the use of AI support for learners as they advance, and critically evaluating the feedback from AI systems instead of authoritatively accepting it because of systematic differences from teacher feedback [8].

8. Limitations

There are some limitations with this review. For Sections 5–6, a narrative/thematic synthesis (rather than meta-analytic) was employed, and strength claims are a qualitative, not a pooled effect size, version. It was limited to English-language publications and therefore could have overlooked certain regional studies and led to a false impression of a geographic focus in East/Southeast Asia. There is limited independent verification of adherence to protocol since the review was not registered on PROSPERO or OSF. Coding judgment calls such as the classification of a region and inclusion of one adjacent science-education study were based on the interpretation of the screening and review teams. Finally, in almost all studies, outcomes have been assessed during or immediately following the intervention, and the durability of gains is untested; the dependency paradox (Section 6.3) has been little tested.

9. Conclusion

Based on 20 studies (2022–2026), the trend is a positive one, and the evidence points to the fact that in most studies, autonomy-related gains are measurable, and they are best understood as a result of three mechanisms in existing autonomy [3], SRL [28], sociocultural [24] and self-determination [21] frameworks. These mechanisms are namely volume-and-availability, low-stakes-practice, and strategy-modeling mechanisms. The persistent "dependency paradox", however, and the lack of follow-up data after the intervention limit the scope of the claims about durable autonomy to that of tool-dependent autonomy. The strongest theme is that AI should be integrated into a "humans-in-the-loop" approach, not replacing human teacher scaffolding; with direct ramifications for teaching and learning, policy and teacher education, and priorities discussed in Section 10.

10. Future Research Directions

Future studies should focus on standardised measures of autonomy, using validated, comparable measures which facilitate meaningful meta-analyses across studies. Longitudinal and post-withdrawal designs are also necessary to decide if the improvements which result from the use of AI are long-term gains in learning or short-

term reliance on technology. Future study of the scaffolding fade-out strategies from the ZPD theory is needed to determine if it is possible to achieve more enduring learning outcomes by gradually fading out the assistance provided by AI than by providing it all at once. For future research, AI–peer and AI–teacher hybrid models that integrate relatedness with autonomy and competence needs can be further studied. Generalizability will be increased by expanding research to contexts outside those of English-language and Asian contexts. Finally, reporting on standardized effect sizes and full statistical information is crucial to further the evidence synthesis.

Creative Commons License Statement

This research work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0>. To view the complete legal code, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode.en>. Under the terms of this license, members of the community may copy, distribute, and transmit the article, provided that proper, prominent, and unambiguous attribution is given to the authors, and the material is not used for commercial purposes or modified in any way. Reuse is only allowed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Conflict of Interest Statement

The author declares no conflicts of interest.

About the Author(s)

Dr. Maha A. Alzahrani is an Assistant Professor of Linguistics at Umm Al-Qura University in Saudi Arabia. Research interests focus on language education, writing skills, and the psychology of English language learners, especially how motivation influences learning. She is also interested in academic literacy and English for Specific Purposes (ESP), exploring how students use language in academic and professional settings.

ORCID: <https://orcid.org/0009-0000-3832-3145>.

References

- 1) Aladini, A., Ismail, S. M., Khasawneh, M. A. S., & Shakibaei, G. (2025). Self-directed writing development across computer/AI-based tasks: Unraveling the traces on L2 writing outcomes, growth mindfulness, and grammatical knowledge. *Computers in Human Behavior Reports*, 17, Article 100566. <https://doi.org/10.1016/j.chbr.2024.100566>
- 2) Apriani, E., Cardoso, L., Obaid, A. J., Muthmainnah, Wijayanti, E., Esmianti, F., & Supardan, D. (2024). Impact of AI-powered chatbots on EFL students' writing

-
- skills, self-efficacy, and self-regulation: A mixed-methods study. *Global Educational Research Review*, 1(2), 57–72. <https://doi.org/10.71380/GERR-08-2024-8>
- 3) Benson, P. (2001). *Teaching and researching autonomy in language learning*. Longman.
 - 4) Chan, S. T. S., Lo, N. P. K., & Wong, A. M. H. (2024). Enhancing university level English proficiency with generative AI: Empirical insights into automated feedback and learning outcomes. *Contemporary Educational Technology*, 16(4), Article ep541. <https://doi.org/10.30935/cedtech/15607>
 - 5) Dizon, G., & Gold, J. (2023). Exploring the effects of Grammarly on EFL students' foreign language anxiety and learner autonomy. *The JALT CALL Journal*, 19(3), 299–316. <https://doi.org/10.29140/jaltcall.v19n3.1049>
 - 6) Fathi, J., & Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: A qualitative study. *Computer Assisted Language Learning*, 39(1–2), 263–302. <https://doi.org/10.1080/09588221.2024.2374772>
 - 7) Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, Article 100055. <https://doi.org/10.1016/j.caeai.2022.100055>
 - 8) Guo, K., & Wang, D. (2024). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29(7), 8435–8463. <https://doi.org/10.1007/s10639-023-12146-0>
 - 9) Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education*, 203, Article 104862. <https://doi.org/10.1016/j.compedu.2023.104862>
 - 10) Han, J., & Li, M. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. *System*, 126, Article 103502. <https://doi.org/10.1016/j.system.2024.103502>
 - 11) Heydarnejad, T. (2025). Unmasking the impacts of self-evaluation in AI-supported writing instruction on EFL learners' emotion regulation, self-competence, motivation, and writing achievement. *Computers and Education: Artificial Intelligence*, 9, Article 100494. <https://doi.org/10.1016/j.caeai.2025.100494>
 - 12) Holec, H. (1981). *Autonomy and foreign language learning*. Pergamon. Retrieved from https://books.google.ro/books/about/Autonomy_and_Foreign_Language_Learning.html?id=NdsWAAAAYAAJ&redir_esc=y
 - 13) Hong, Q. N., Pluye, P., 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). *Mixed Methods Appraisal Tool (MMAT), version 2018*. Registration of Copyright (#1148552), Canadian Intellectual Property Office, Industry Canada.
 - 14) Jmaiel, H. A., Abukhait, R. O., Mohamed, A. M., Shaaban, T. S., Al-khresheh, M. H., & AL-Qadri, A. H. (2025). The role of ChatGPT in enhancing EFL students' ESP writing skills: An experimental study of gender and major differences. *Discover Education*, 4, Article 240. <https://doi.org/10.1007/s44217-025-00700-6>

- 15) Little, D. (1991). *Learner autonomy: Definitions, issues and problems*. Authentik. Retrieved from https://www.researchgate.net/publication/259874253_Learner_Autonomy_1_Definitions_Issues_and_Problems
- 16) Liu, Z. M., Hwang, G. J., Chen, C. Q., Chen, X. D., & Ye, X. D. (2024). Integrating large language models into EFL writing instruction: Effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 39(3), 466–490. <https://doi.org/10.1080/09588221.2024.2389923>
- 17) Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11, Article 9. <https://doi.org/10.1186/s40561-024-00295-9>
- 18) Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, 55(4), 1328–1353. <https://doi.org/10.1111/bjet.13454>
- 19) Nguyen, Q., & Doan Thi Hue, D. (2025). ChatGPT's effects on EFL learners' autonomy and self-efficacy in self-regulated learning contexts. *The JALT CALL Journal*, 21(2), Article 102968. <https://doi.org/10.29140/jaltcall.v21n2.102968>
- 20) Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
- 21) Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. Retrieved from https://selfdeterminationtheory.org/SDT/documents/2000_RyanDeci_SDT.pdf
- 22) Shi, L., Li, S., & Xing, J. (2025). Exploring Chinese secondary EFL students' self-regulated learning and task engagement in AI-assisted classrooms: A latent growth curve modelling study. *European Journal of Education*, 60(4), Article e70241. <https://doi.org/10.1111/ejed.70241>
- 23) Teng, M. F., & Huang, J. (2026). Assessing ChatGPT feedback for EFL learners' engagement and writing performance. *International Journal of Applied Linguistics*. Advance online publication. <https://doi.org/10.1111/ijal.70155>
- 24) Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. Retrieved from https://books.google.ro/books/about/Mind_in_Society.html?id=RxjjUefze_oC&redir_esc=y
- 25) Wang, Y. (2025). Reducing anxiety, promoting enjoyment and enhancing overall English proficiency: The impact of AI-assisted language learning in Chinese EFL

- contexts. *British Educational Research Journal*. Advance online publication. <https://doi.org/10.1002/berj.4187>
- 26) Wang, Y. (2026). Enhancing university EFL learners' writing performance: The role of AI-enhanced goal-setting, feedback and social norm interventions. *Journal of Computer Assisted Learning*, 42(1), Article e70179. <https://doi.org/10.1002/jcal.70179>
- 27) Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. *Frontiers in Psychology*, 14, Article 1249991. <https://doi.org/10.3389/fpsyg.2023.1249991>
- 28) 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. Retrieved from https://books.google.ro/books/about/Handbook_of_Self_Regulation.html?id=rv3DZSim6z4C&redir_esc=y

Appendices

Appendix A: Search Strategy Log

Table A1 reports the database, search date, and records retrieved using the combined search string in Section 4.3, adapted to each database's syntax.

Table A1: Search strategy log by database and source

Database / Source	Date Searched	Search Field(s)	Records Retrieved
ERIC	March 3, 2026	Descriptor / Abstract	312
Scopus	March 3, 2026	TITLE-ABS-KEY	486
Web of Science	March 4, 2026	Topic	401
ScienceDirect	March 4, 2026	Title, abstract, keywords	227
Taylor & Francis Online	March 5, 2026	Full text (title/abstract limited)	184
Wiley Online Library	March 5, 2026	Title, abstract, keywords	232
Google Scholar (grey literature)	March 6–7, 2026	All fields (first 10 pages of results)	41
Backward/forward citation chasing	March 8–10, 2026	Reference lists and citing articles of initially screened records	22
Total records identified	—	—	1,905

Appendix B: Screening Flow Record

Table B1 reports records retained and excluded at each PRISMA screening stage, corresponding to Figure 2.

Table B1: Screening flow record with exclusion reasons by stage

Stage	n Assessed	n Excluded	Primary Exclusion Reasons
Records after duplicate removal	1,293	—	612 duplicates removed from 1,905 identified records
Title/abstract screening	1,293	1,180	Clearly non-EFL, non-writing, or non-AI focus; non-empirical format apparent from title/abstract
Full-text eligibility assessment	113	93	Population/context not EFL (28); no writing-skill focus (14); non-empirical (22); outcome not autonomy/SRL/engagement-related (17); not in English or unverifiable outlet (7); duplicate/overlapping sample (5)
Studies included in synthesis	20	0	Met all PICOS and Table 2 criteria

Appendix C: Data Extraction Matrix

Table C1: Summary of Included Studies and Autonomy-Related Findings

Author(s), Year	Context	N	Design	AI Tool Type	Autonomy-Related Finding
[5]	Japan, university	-	Pre/post, single group	AWE (Grammarly)	Significant decrease in FLA and increase in learner autonomy
[19]	Vietnam, university	50	RCT (mixed-methods)	Generative (ChatGPT)	Greater autonomy gains than teacher-led group; equal self-efficacy gains
[16]	China, elementary	65	RCT	Generative (LLM-CALLA)	Increased performance, SRL strategy use, motivation; sustained at follow-up
[6]	Not specified, tertiary	-	Qualitative	Generative (AI portfolio)	Increased emotion regulation, mindfulness, confidence
[7]	Japan, adult	10	Within-subject comparison	Custom (AI KAKU)	Comparative outcomes not confirmed
[8]	China, university	50 essays	Quasi-experimental/comparative	Generative (ChatGPT)	Greater, more evenly distributed feedback vs. teachers
[23]	China, university	174	Quasi-experimental	Generative (ChatGPT)	Increased affective/behavioral engagement; increased writing performance
[10]	China, university	-	Mixed-methods	Generative + hybrid model	Increased engagement with AI-supported revision
[17]	Not specified, tertiary	-	Mixed-methods	Generative (ChatGPT)	Significant increase in academic writing skills
[9]	Not specified	-	Comparative intervention	Custom (chatbot debate)	Increased argumentation quality, enjoyment, effort
[25]	China	80	Mixed-methods	Custom (AI app)	Decreased anxiety, increased enjoyment
[1]	Iran (Ahvaz)	561	Quasi-experimental	Custom (self-directed AI tasks)	Statistical outcomes not confirmed

Maha A. Alzahrani
FOSTERING LEARNER AUTONOMY IN EFL WRITING
THROUGH ARTIFICIAL INTELLIGENCE: A SYSTEMATIC REVIEW

[22]	China, secondary	334	Latent growth curve model	Custom (AI-assisted classroom)	Modest but significant SRL/engagement growth; ceiling effect for high-SRL learners
[4]	Hong Kong, university	918	RCT	Generative (GPT-3.5)	Increased essay quality on revision; increased engagement
[11]	Not specified	-	Quasi-experimental	Custom (AI-supported self-evaluation)	Increased emotion regulation, self-competence, motivation, achievement
[26]	Not specified, university	383	Four-arm RCT	Custom (AI nudges)	Large, significant post-test differences; goal-setting nudge strongest
[14]	Not specified, university	117	Experimental	Generative (ChatGPT, ESP)	Substantial writing gains; larger for female students
[2]	Indonesia (IAIN Curup)	40	Mixed-methods, quasi-experimental	Generative (AI chatbot)	Increased proficiency, self-efficacy, self-regulation; overreliance concerns noted
[18]	Hong Kong, secondary	74	3-week pilot comparison	Custom (SRLbot vs. Nemobot)	Generative chatbot outperformed rule-based chatbot on SRL (science-ed context)
[27]	China, university	190	RCT, repeated measures	AWE (Grammarly)	Consistent gains across all IELTS writing dimensions

Appendix D: Thematic Codebook

Table D1 documents the definition and an illustrative example for each of the six themes in Section 5.3.

Table D1: Thematic codebook with theme definitions and illustrative examples

Theme	Definition	Illustrative Example
1. AI feedback and autonomy construction	Findings describing how the volume, focus, or distribution of AI-generated feedback shapes learners' independent engagement with revision	[8]: ChatGPT feedback more evenly distributed across content, organization, and language than teacher feedback
2. AI-supported SRL strategy construction	Findings describing AI features that model or scaffold specific self-regulated learning strategies (goal-setting, monitoring, evaluation)	[26]: goal-setting nudge produced the strongest post-test outcome among three AI-enhanced interventions
3. Autonomy and self-efficacy outcomes	Findings reporting change on validated or study-specific autonomy and self-efficacy measures	[19]: significantly greater autonomy gains for ChatGPT-facilitated vs. teacher-facilitated group
4. Affective and motivational dimensions	Findings concerning anxiety, enjoyment, motivation, or emotion regulation associated with AI use	[5]: significant reduction in foreign language anxiety with sustained Grammarly use
5. Writing performance outcomes	Findings reporting change in writing quality, proficiency, or test-based writing scores	[27]: consistent gains across all IELTS writing dimensions following AWE training
6. Tensions and countervailing evidence	Findings identifying risks, ceiling effects, or unconfirmed/mixed outcomes that qualify the otherwise positive evidence base	[22]: ceiling effect suggesting diminishing returns for already highly self-regulated learners