

Generative AI and Adaptive Systems for Customising Help-Seeking Scaffolds: A Systematic Review

Jecha Jecha

jechasulieman@gmail.com

Southwest University <https://orcid.org/0000-0002-7784-0994>

Chimaobi Charles Igwean

Southwest University

Anum Banaras

Southwest University

Systematic Review

Keywords: Generative artificial intelligence, Adaptive learning systems, Help-seeking scaffolds, Self-regulated learning, Personalised feedback

Posted Date: April 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9471902/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Help-seeking is a central self-regulated learning (SRL) process; however, generic scaffolds often misalign with learners' heterogeneous needs. This systematic review addresses the lack of consolidated evidence on how artificial intelligence (AI), including large language models (LLMs), customises help-seeking scaffolds, the effects thereof, and the methodological and ethical constraints. Guided by SRL theory, this review examined empirical studies on AI-driven customised help-seeking support, focusing on evidence availability, comparative impacts versus standard scaffolds, evaluation approaches, and reported gaps. Database searches (IEEE Xplore, Web of Science, ERIC, and ScienceDirect) conducted on 28 May 2025 yielded 3,576 records; blinded multi-reviewer screening in Rayyan following PRISMA procedures identified 31 eligible studies for narrative synthesis. Data extraction by three reviewers showed high raw agreement (87.3%) but low chance-corrected agreement (Cohen's $\kappa \approx 0.05$; Fleiss' $\kappa \approx 0.17$), resolved via structured adjudication and full consensus. The included interventions (2017–2025) ranged from rule-based intelligent tutoring and learning-analytics dashboards to LLM-based textbooks, writing feedback, laboratory guidance, and coding assistants, enabling conversational, intent-sensitive, on-demand scaffolding. Most studies reported improved academic performance and engagement, particularly for lower-performing or lower-prior knowledge learners, although null or mixed effects and inconsistent satisfaction outcomes were common. The recurrent limitations concerned sample scope, duration, measurement heterogeneity, usability, and risks related to hallucinations, bias, privacy, equity, and overreliance. The review concludes that although AI-customised help-seeking scaffolds are promising, they are context-dependent and delineate methodological and design priorities to advance SRL-aligned, responsible AI support for help-seeking.

Introduction

Help-seeking is a well-established self-regulated learning (SRL) strategy: learners who recognise difficulties and actively seek assistance tend to outperform those who do not (Karabenick & Newman, 2006). Instructional scaffolds hints, prompts, and models delivered at the right moments improve learning outcomes by building competence within learners' zone of proximal development until support can be withdrawn (Puntambekar & Hübscher, 2005; Reiser, 2004).

Adaptive learning systems intelligent tutoring systems, learning analytics dashboards, and AI-driven platforms have long addressed this need by monitoring learner behaviours and adjusting feedback accordingly (Gomes, 2024; Roll & Winne, 2015). The latest wave of educational technology extends these capabilities through generative AI and large language models (LLMs) such as ChatGPT, which can engage in open-ended dialogue, interpret the intent behind a query, and generate varied explanations in real time (Kasneci

et al., 2023). This shift from rule-based adaptation to conversational co-construction of support represents a qualitatively different form of scaffolding: on-demand, intent-sensitive, and unconstrained by predefined scripts (Hong et al., 2025; Naseer & Khawaja, 2025).

Despite rapid adoption, rigorous empirical evidence on how these tools scaffold help-seeking remains limited (Cheong et al., 2024; Zaharuddin et al., 2024). There is little systematic research comparing AI-mediated help-seeking with conventional scaffolds, and the effectiveness of AI-generated explanations for skill development is unclear (Lai, 2024). ChatGPT compared with human experts or conventional systems (Bit et al., 2024). Similarly, although generative AI can produce personalised explanations, the effectiveness of these AI-generated scaffolds in improving understanding and skill development remains unclear. Overall, the application of generative AI in higher education is underexplored and its educational impact is not well documented (Lai, 2024; Mcgrath et al., 2024).

This systematic review addresses that gap. By synthesising empirical evidence through the lens of SRL theory, this review identifies cross-study patterns, evaluates outcomes, and assesses design strategies and persistent challenges (Guettala et al., 2024; Xiao et al., 2025).

Zimmerman's (2002) SRL framework guides this review because it positions help-seeking as a fundamental regulatory strategy: self-regulated learners recognise difficulties and seek assistance rather than persisting in confusion or abandoning a task (Karabenick & Newman, 2006; Panadero, 2017). This lens allows us to evaluate whether AI-driven scaffolds facilitate productive help-seeking processes, not just performance gains (Kong & Yang, 2024). The review addresses four research questions: Instead of persisting in confusion or abandoning a task, a self-regulated learner recognises the need for assistance and promptly seeks it (Karabenick & Newman, 2006; Memon et al., 2022). The SRL perspective is valuable for our review because it focuses on the learning process (not just outcomes) and places help-seeking within a wider range of regulatory behaviours (Panadero, 2017). By adopting SRL as the theoretical foundation, we can explore whether and how generative AI and adaptive systems facilitate productive help-seeking processes (Kong & Yang, 2024; Sardi et al., 2025).

1.1 Research questions

1. What evidence exists regarding the use of generative AI and adaptive systems to customise help-seeking scaffolds for learners in educational settings?
2. Compared to standard or non-customised scaffolds, how do AI-driven customised scaffolds impact learner engagement, academic performance, and satisfaction?
3. What are the standard methodologies and study designs used to evaluate the effectiveness of generative AI and adaptive systems in scaffold customisation?

4. What gaps and limitations have been reported in studies on AI-driven help-seeking scaffold customisation in education?

2 Literature Review

2.1 Importance of Help-Seeking and Scaffolding in Education

Effective learning requires students to recognise when they need help and seek it strategically. This help-seeking behaviour is central to self-regulated learning, yet students often struggle to identify appropriate help sources.

Instructional scaffolding, rooted in Vygotsky theory, provides strategic support that is gradually withdrawn as learner competence increases. In help-seeking contexts, scaffolds enable learners to recognise when they need help and seek it effectively. Traditional scaffolding relied on teachers to diagnose gaps and provide support; scaling this remains challenging with large heterogeneous classes.

2.2 Evolution of Traditional and Adaptive Scaffolding Systems

Scaffolding has evolved from static, teacher-delivered support to dynamic, technology-improved approaches that adapt to learner needs.

Soft scaffolding emerged to provide dynamic, responsive support that adapts to learner needs, addressing limitations of static systems. It provides support on the appropriate level of support needed, similar to how a human tutor personalises instruction (Ginting et al., 2024; Singh et al., 2024; Walker et al., 2011). Research in this field has shown that adaptive scaffolding tends to support self-regulated learning more effectively than fixed support because it can be adjusted to meet students' immediate learning needs (Roll et al., 2014). Students who receive adaptive help are guided by obstacles with timely feedback, ideally without revealing the solution, thus maintaining an optimal challenge level.

Traditional adaptive systems (e.g. rule-based ITS) have limitations: they require extensive manual rule definition, cannot easily adapt to novel domains, and provide static support. These constraints motivated researchers to explore AI-based approaches.

2.3 Emergence of Generative AI in Personalised Learning Support

Recent AI advancements, particularly large language models, enable powerful personalisation of help-seeking support. LLMs can generate context-aware, responsive feedback at scale, addressing limitations of earlier adaptive systems.

LLM-based support can simulate a "more knowledgeable other", providing contextual scaffolding that adapts to learner needs. This dynamic, responsive support represents a qualitative shift from earlier adaptive systems.

The capabilities of generative AI offer promising solutions to the persistent limitations of adaptive scaffolding. Notably, they significantly alleviate the authoring burden, as educators no longer need to pre-program every hint; AI can generate novel assistance in real time (Rizvi, 2023). This adaptability also allows scaffolds to be highly context-aware. A single AI model, if adequately trained and guided by effective prompts, can potentially support learning across various domains by leveraging extensive knowledge. Early evidence of the efficacy of AI-based scaffolding is promising (Lee et al., 2022; Siu et al., 2025). Students who used ChatGPT or similar AI tools often demonstrate improved learning performance and engagement levels. For example, a recent meta-analysis of studies from late 2022 to early 2025 found that using the ChatGPT had a substantial positive effect on students' learning outcomes. These results suggest that when used appropriately, generative AI can improve certain aspects of learning by providing timely and personalised content (Ali et al., 2024; Anjum et al., 2025). Additionally, AI-driven scaffolds can assume multiple roles, such as an intelligent tutor offering explanations, a learning companion prompting reflection, or a brainstorming tool, making them versatile additions to educational environments (Rathika et al., 2024).

However, integrating generative AI into learning support presents challenges, particularly regarding ethical implications and content quality. Key concerns include data privacy, potential AI-generated inaccuracies, and equitable access. These limitations must be addressed in research and deployment.

Additionally, this review aims to guide future research by identifying methodological and evaluative shortcomings in existing studies and proposing directions for advancing the reliability assessment of AI-based scaffolding. Drawing on ethical AI frameworks and prior work on bias in AI-improved education (Cheng et al., 2015), this review suggests the need for new metrics and evaluative frameworks, such as accuracy confidence indices, epistemic reliability markers, and culturally responsive validity checks, to systematically assess the quality of AI-generated scaffolds. These measures can inform future standards for the responsible integration of generative AI into educational settings, so that scaffolding remains pedagogically sound and ethically grounded.

2.4 Gaps in Research and Need for Synthesis

AI-improved help-seeking scaffolds represent a rapidly developing research frontier. This review consolidates disparate evidence to uncover patterns and identify research gaps.

By consolidating these disparate strands of evidence, this review seeks to uncover novel patterns and theoretical linkages that clarify how AI-driven scaffolds support or hinder SRL phases, including goal-setting, strategy use, and self-monitoring (Faza & Lestari, 2025). In doing so, it advances the conceptual framework for effective design patterns specific to AI-improved help-seeking, moving beyond the generic recommendations typically found in general syntheses of educational technology. This integrative approach identifies empirical consistencies and divergences and maps how different forms of AI scaffolding (e.g. adaptive hints, generative feedback, or conversational tutoring) align with SRL constructs, thereby deepening theoretical understanding and guiding pedagogical design (Lan & Zhou, 2025).

Another issue compounded by study fragmentation is the inconsistent use of measurement tools to evaluate help-seeking and SRL outcomes. Current studies employ heterogeneous instruments ranging from self-report surveys and log data to performance scores, making cross-study comparisons difficult. Building on comprehensive reviews of AI-improved personalised learning systems (Guettala et al., 2024), this review proposes the need for a standardised evaluative framework that incorporates multidimensional metrics. Such a framework could include (a) behavioural indicators of adaptive help-seeking (e.g. frequency, timing, and relevance of help requests), (b) cognitive measures of metacognitive awareness and regulation, and (c) affective dimensions, such as learner confidence and trust in AI feedback. Establishing consistent metrics of this kind would enable a more robust meta-analytical synthesis, but also provide a foundation for assessing the educational quality and SRL relevance of AI-generated scaffolds in future research.

Quality and safety concerns around AI-generated scaffolding content remain largely unexplored in the literature. Most studies lack explicit quality checks or content validation mechanisms.

2.5 Theoretical Framework

Before presenting the theoretical framework, it is necessary to clarify the conceptual boundaries of this review. We define a *help-seeking scaffold* as an instructional support mechanism that (a) is activated in response to a learner-initiated request for assistance or a system-detected learning impasse, and (b) provides graduated, context-sensitive guidance aimed at resolving the impasse while promoting the learner's capacity for independent problem-solving. This definition distinguishes help-seeking scaffolds from broader forms of AI-supported personalised learning, such as content sequencing based on learning-style inventories, automated grammar correction, or AI-generated instructional media, which deliver adaptive support but do not specifically scaffold the help-seeking process. Accordingly, included studies were classified as either *core* (directly meeting both criteria) or *adjacent* (providing adaptive learning support that partially overlaps with

help-seeking but does not fully meet the operational definition). Both categories are retained in the synthesis to preserve comprehensiveness, but findings are reported with this distinction to ensure analytical precision (see Section 3 and Appendix A).

2.5.1 SRL served as the theoretical foundation for this study

Self-Regulated Learning (SRL) served as the foundation for this review. SRL comprises forethought (planning), performance (execution), and reflection, explaining how learners manage cognition, metacognition, motivation, and behaviour. Help-seeking is a critical metacognitive process within SRL: learners must identify knowledge gaps, decide when to seek help, engage with support, and reflect on effectiveness. AI-customised help-seeking scaffolds improve learners' self-regulatory capacity by supporting goal-setting, progress monitoring, and strategy reflection.

2.5.2 SRL-Informed Design and Evaluation of Scaffolds

Building on SRL theory has significant implications for designing and evaluating help-seeking scaffolds. Modern SRL-informed scaffolds are meticulously designed to target specific regulatory processes and encourage learners to adopt the strategies employed by expert self-regulators. Instead of merely providing answers, SRL-aligned scaffolds promote goal setting ("What is your plan for tackling this problem?"); monitoring ("Have you checked your work for errors?"), and strategic planning ("What strategy could you try next?") learning process (Azevedo et al., 2022; Andrade, 2019; Lim et al., 2022). When students encounter difficulties, SRL-based help-seeking scaffolds offer hints that encourage such reflections. The aim is to improve, rather than circumvent, learners' self-regulation, thereby enabling them to internalise effective learning habits over time (Järvelä et al., 2023).

In practice, educational technologies incorporate SRL scaffolding through self-checklists, metacognitive prompts, and pedagogical agents that act as external regulators (Y. Guo et al., 2023). For example, the MetaTutor system uses multiple agents to prompt students to plan, monitor their understanding, and reflect on complex science learning tasks (Azevedo et al., 2019). These agents intervene with prompts based on SRL models and are designed to shift regulatory responsibility from software to students.

Evaluating SRL-informed scaffolds requires moving beyond test scores to include metacognitive awareness, self-efficacy, and help-seeking behaviour. Comprehensive evaluation aligns theory with practice.

2.5.3 Justification for an SRL Conceptual Foundation

Grounding this systematic review in Self-Regulated Learning (SRL) theory was a deliberate choice because of the strong alignment between the goals of help-seeking scaffolds and the core principles of SRL. Both scaffolding and SRL aim to empower students to become autonomous and effective learners by transitioning from assisted performance to independent competence (Miao and Ma, 2023). SRL theory provides a well-researched framework for independent learning, emphasising goal-directedness, self-monitoring, strategies, and reflection. Anchoring this review in SRL ensures that AI-driven and adaptive scaffolding tools are evaluated for both their immediate performance and their role in supporting students' self-regulatory skills (Taub et al., 2020).

SRL captures the complexity of help-seeking by integrating cognitive, metacognitive, motivational, and social factors.

By examining generative AI and adaptive systems as scaffolds for SRL, this review is positioned at a critical intersection of research and practice, exploring how these technologies can support students' capacities for autonomous help-seeking and lifelong learning (Banihashem et al., 2025; Faza & Lestari, 2025).

The model in Fig. 1 integrates SRL with help-seeking through cognitive/metacognitive, motivational/affective, and social-interactive factors. This framework guides design and evaluation of adaptive scaffolding systems.

Three categories of influence—cognitive/metacognitive, motivational/affective, and social-interactive—were systematically mapped onto all phases. Cognitive/metacognitive factors include awareness of task demands and knowledge of strategies used. Motivational and affective constructs encompass self-efficacy, goal orientation, and perceived value of help-seeking. Socio-interactive aspects reflect the environmental conditions and norms that shape learners' willingness to seek help from peers. The structure of the diagram demonstrates that help-seeking acts as a contingent intervention when self-regulation stalls (impasses) and as a process influenced by broader motivational and social dynamics. Thus, it does not bypass self-regulation but reinforces the cycle. The outcome box, Improved SRL," emphasises that successful scaffolding and help-seeking should ultimately support greater learner autonomy and strategic competence.

This integrated SRL-help-seeking model offers a nuanced framework for designing adaptive scaffolding and evaluating interventions. Measurements should extend beyond performance metrics to capture shifts in self-regulatory behaviours (e.g. increased planning, calibrated self-assessment, and strategic help use), as advocated by

contemporary SRL studies. Thus, the figure operationalises the core premise of SRL theory: scaffolds and interventions should support transferable regulatory skills, equipping learners with independent task management over time. In summary, SRL theory provides a principled and contemporary foundation that keeps the review focused on the long-term development of self-reliant, reflective, and adaptive learners, so that the excitement around generative AI and adaptivity is always connected to deep learning goals.

Methodology

3.1 Search strategy and study selection

The original search was conducted on 28 May 2025 across four databases (IEEE Xplore, Web of Science, ERIC, and ScienceDirect) that were exported in the RIS format based on database-based identification of publications on technology-improved learning reviews. The final database with references was converted to Rayyan, enabling blinded independent decisions, conflict resolution, and documentation of exclusion reasons (Abbas et al., 2022; Hernandez-Cruz et al., 2024; Ouzzani et al., 2016). The searches were created to capture the most recent empirical literature on generative AI and help-seeking scaffolds for higher education and were updated through 2025. The search phrases used were *(learner OR student OR universe* OR higher education) AND (generative AI OR gen-AI OR adaptive learning OR personalised learning OR intelligent tutor*) AND (support* OR scaffold* OR intervention*) AND (academic performance OR learning outcome*)*, in line with the current terminology in contemporary Gen AI-in-education research.

The study selection process was conducted in two phases.

In Phase 1, three reviewers examined each citation and abstract to assess eligibility using Rayyan's functionality for blinding and conflict marking. This process allowed for the registration of independent decisions and conflicts, which were resolved through discussion and, when necessary, by consulting the full texts (Valizadeh et al., 2022).

In the second stage, the same three reviewers independently read all candidate papers and recorded the reasons for exclusion (including the absence of a generative AI component, help-seeking, scaffolding in AI, and adaptive systems). This resulted in 31 studies that referred to one or more AI and adaptive systems and help-seeking outcomes of the use of generative AI, which constituted sample-relevant qualitative syntheses. We also followed the updated PRISMA guidelines for the transparent reporting of study selection (Page et al., 2021). We present a PRISMA-style flow diagram to summarise the identification, screening, eligibility, and inclusion steps of this study.

3.2 Inclusion and Exclusion Criteria

Articles were considered based on the inclusion and exclusion criteria listed in Table 1. A database of 3576 articles containing titles, abstracts, and full texts was created using Rayyan reference management software. The results are summarised in Appendix A. Figure 2 provides an overview of our search protocol based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (Page et al., 2021).

Table 1
Inclusion and Exclusion Criteria

Inclusion	Exclusion
Studies published in peer-reviewed journals or conference proceedings	Non-peer-reviewed articles, opinion pieces, editorials.
Research focused on generative AI and/or adaptive systems applied to customising help-seeking scaffolds in educational settings.	Studies not focused on help-seeking scaffolds or customisation.
Empirical studies (quantitative, qualitative, or mixed methods).	Research outside educational contexts (e.g., healthcare, unless explicitly about learning).
Studies involving learners (K-12, higher education, adult learning).	Theoretical papers without empirical data, review articles, and opinion pieces without systematic methods.
Publications in English	Non-English study

Following study selection, each of the 31 included studies was classified as either *core* or *adjacent* based on the operational definition of help-seeking scaffolds presented in Section 2.5. Core studies (n = 25) investigated interventions that (a) were activated by learner-initiated requests or system-detected impasses and (b) provided graduated guidance to resolve those impasses. Adjacent studies (n = 6) examined AI-driven adaptive learning supports that partially overlapped with help-seeking (e.g. learning-style-based content sequencing, AI-generated instructional videos, automated grammar correction tools, or game-based incentive mechanisms) but did not directly scaffold the help-seeking process. Two reviewers independently classified each study, with disagreements resolved through discussion. Adjacent studies were retained in the synthesis to capture the wider adaptive learning context but are clearly distinguished in the results to maintain analytical focus on help-seeking scaffolds (see Appendix A for the classification of each study).

3.3 Data Extraction

Data were extracted using a structured form in Microsoft Excel, following established protocols for systematic reviews (Hill et al., 2023). A standardised extraction spreadsheet was created to ensure clarity and consistency in the team. Three reviewers independently extracted data from each of the 31 included studies and compiled an Excel spreadsheet

with predefined data fields. The extracted key information included Study Identification and Context (citation details, year, and educational setting), Intervention Details (type of AI or adaptive technology, scaffold design, and customisation approach), Study Design and Sample (methodology, sample size, and participant characteristics), Outcomes and Measures (academic, engagement, and related indicators, including measurement approaches and intervention duration), Key Findings, and Limitations (biases, sample size, and technical constraints) (Ayeni et al., 2024).

To ensure reliability, all reviewers independently completed the extraction process, and their results were cross-checked for consistency. Inter-rater reliability (IRR) was calculated using Cohen's kappa coefficient. This level of concordance aligns with the established thresholds for excellent reliability in educational and psychological research (Azevedo et al., 2022; T. Li et al., 2023). Any remaining discrepancies were discussed and resolved through consensus, following the best practice verification procedures for systematic reviews. This rigorous double-checking process ensured the accuracy and reproducibility of the extracted data, with a harmonised extraction table serving as the foundation for qualitative synthesis (Wu et al., 2024).

3.4 Inter-Coder Reliability

The team evaluated inter-coder reliability at each major review stage (McHugh, 2012). During title/abstract screening, agreement among the three independent reviewers was assessed using pairwise Cohen's kappa coefficients and Fleiss' kappa for overall three-reviewer agreement, with the latter accounting for multiple raters and chance agreement (Moons & Vandervieren, 2025). Raw agreement was high (87.3%); however, chance-corrected kappa values were low (Cohen's $\kappa \approx 0.05$; Fleiss' $\kappa \approx 0.17$), which indicates only slight agreement beyond chance. This discrepancy is attributable to the prevalence effect: because both reviewers excluded the vast majority of references, the marginal distributions were highly skewed, which inflates raw agreement while depressing kappa (Byrt et al., 1993; Cicchetti & Feinstein, 1990; Vanbelle & Albert, 2009). To account for this, the prevalence-adjusted bias-adjusted kappa (PABAK) was calculated at 0.746, indicating substantial agreement once prevalence and bias are corrected (Byrt et al., 1993). Nevertheless, the low standard kappa values represent a methodological limitation that is fully acknowledged (see Section 4.1). All 453 discordant cases (12.7%) were independently re-evaluated by a third reviewer, and disagreements were resolved through structured discussion until full consensus was reached. Inter-coder reliability was also addressed during data extraction and quality appraisal. The extraction form was pilot-tested and refined to minimise interpretive divergence, and all three reviewers independently extracted data as previously described (Bertizzolo et al., 2019). The extracted entries were compared, and disagreements requiring consensus discussions were infrequent. For quality appraisal, calibration on example studies preceded

independent ratings, and any rating discrepancies were resolved through discussion until consensus was reached. These steps ensured consistent inclusion criteria, uniform data recording, and reliable evaluation of study quality across all review stages (Nussbaumer-Streit et al., 2023).

3.5 Quality Appraisal

In our systematic review, we conducted a quality appraisal of each included study to assess its methodological rigor and risk of bias. To ensure transparency and reproducibility, the studies were rated as *High*, *Moderate*, or *Low* according to the number of methodological domains that met the established appraisal criteria (Azarian et al., 2023). The evaluation considered five key domains adapted from the Mixed Methods Appraisal Tool (MMAT) (Hong et al., 2022), which is designed for systematic reviews that include studies with diverse designs (quantitative, qualitative, and mixed methods). These five domains, mapped from the MMAT criteria for quantitative and mixed-methods studies, included (1) robustness of study design and comparability between groups, (2) internal validity reflected through the control of confounders and intervention fidelity, (3) adequacy and representativeness of the sample, (4) validity and consistency of data collection and measurement approaches, and (5) completeness and transparency of reporting (Chen et al., 2023; Hong et al., 2022). Each domain was assessed as either meeting or not meeting quality criteria. Studies meeting four or five domains were rated as *high quality*, which indicates strong methodological rigor and minimal bias. Studies meeting two or three domains were considered *to be of moderate quality*, indicating some limitations but acceptable validity, whereas studies meeting only one or none of the domains were classified as *low quality*, which indicates significant methodological weaknesses or insufficient reporting (Wilke et al., 2015).

Three reviewers independently applied these criteria across all 31 studies and provided brief justifications noting notable strengths and limitations, such as rigorous experimental design or unaddressed confounding variables. The reviewers' assessments generally converged, and any discrepancies were resolved through detailed discussions until a consensus was reached. No study was excluded solely based on quality, which is consistent with the best practices for systematic synthesis. Instead, ratings were used to guide the interpretation and weighting of evidence, with more cautious conclusions drawn from lower quality studies. This transparent rubric-based approach ensured that methodological rigor was consistently applied across the corpus and that the synthesis accurately reflected the strength of the available evidence.

3.6 Data Synthesis

We employed a narrative synthesis approach to compile the evidence from the included studies. No meta-analysis was conducted; the heterogeneity of interventions, contexts, and

outcome measures rendered statistical pooling inappropriate. studies, rendering statistical pooling inappropriate (He et al., 2022). In line with contemporary guidance for narrative synthesis in systematic reviews (Hong et al., 2022), we used textual descriptions to systematically summarise the findings from diverse studies. The extracted data were organised according to the key themes and focal aspects of the research questions. Studies were categorised according to the type of adaptive learning technology or scaffold used to facilitate outcome comparisons (Siddaway et al., 2019).

We developed descriptive summaries showing patterns, agreements, and divergences across studies. Studies were coded against Zimmerman's SRL model to systematically ground the typology.

To systematically ground the SRL-informed typology, each included study was coded against Zimmerman's (2002) three-phase SRL model. Three phases were operationalised as follows: (1) *Forethought* the scaffold prompts goal-setting, task analysis, or strategic planning before task execution (for example, "What is your plan?" prompts or pre-task goal-setting interfaces); (2) *Performance* the scaffold provides real-time feedback, error correction, hints upon impasse detection, or strategy adjustment during task execution (e.g. adaptive hints during problem-solving, ChatGPT explanations during lab work); and (3) *Self-Reflection* the scaffold prompts evaluation of outcomes, attribution analysis, or strategy revision after task completion (e.g. post-task reflection prompts, revision feedback on essays). A study could be coded to multiple phases if its intervention addressed more than one. Two reviewers independently coded each study, with disagreements resolved through discussion. The resulting SRL phase coding for each study is reported in Appendix A and summarised in the cross-tabulation presented in the Results (Table 7).

Results

4.1 Screening Reliability

Two independent reviewers screened 3,576 unique references based on their titles and abstracts. The reviewers agreed to 87.3% of the inclusion and exclusion criteria for the study. In 453 (12.7%) cases, the decisions were discordant and required assessment by a third reviewer. All conflicts of interest were resolved by consensus among the authors. Table 2 summarises the inter-rater reliability (IRR) outcomes for the screening process, including raw agreement and chance-corrected agreement statistics (Cohen's κ and Fleiss' κ), as recommended for systematic reviews (Wang et al., 2020; Yao et al., 2024).

Table 2
Inter-Rater Reliability Results.

Measure	Result
References screened (title/abstract)	3,576
Initial conflicts (disagreements)	453 (12.7% of references)
Raw agreement (two reviewers)	87.3%
Cohen's κ (two reviewers)	~ 0.05 (Slight agreement)
Fleiss' κ (three reviewers)	~ 0.17 (Slight agreement)
PABAK (two reviewers)	0.746 (Substantial agreement)
<i>Note: Cohen's κ was calculated between the two primary reviewers; Fleiss' κ included the third reviewer. PABAK (prevalence-adjusted bias-adjusted kappa) corrects for the skewed marginal distributions caused by high exclusion prevalence (Byrt et al., 1993).</i>	

Table 2 shows that the reviewers demonstrated 87.3% raw agreement, indicating consistent screening judgments for most references. However, the chance-corrected statistics were markedly lower: Cohen's κ was approximately 0.05 and Fleiss' κ approximately 0.17, with both indicating *slight agreement* beyond chance (Hirt et al., 2021). This discrepancy between high raw agreement and low kappa is a well-documented statistical artefact known as the “prevalence effect” or “paradox of kappa”: when the trait being coded is rare (here, inclusion), marginal distributions become highly skewed, inflating observed agreement while depressing kappa (Byrt et al., 1993; Cicchetti & Feinstein, 1990). To quantify this, the prevalence-adjusted bias-adjusted kappa (PABAK) was calculated: $PABAK = 2(0.873) - 1 = 0.746$, which falls in the substantial agreement range (Byrt et al., 1993). Nevertheless, the low standard kappa values represent a genuine methodological limitation of the initial screening phase and warrant explicit acknowledgement (Guyatt et al., 2025).

A structured, multistage consensus process was implemented to mitigate the potential impact of this limitation. After the initial screening, all 453 discordant cases were independently re-evaluated by a third reviewer using predefined inclusion and exclusion criteria. Rather than serving as a simple tiebreaker, the third reviewer facilitated deeper, criteria-based discussions among all authors. This re-evaluation involved revisiting each borderline case to ensure that inclusion decisions were justified by clear evidence of alignment and methodological fit, in accordance with a systematic review of best practices for managing low κ -values (Braithwaite et al., 2019; Lipman et al., 2023). Through this rigorous adjudication process, 100% consensus was achieved, and the final study selection was verified against the full eligibility criteria to eliminate residual bias. This approach ensured that, despite the low chance-corrected agreement, the final corpus of the included studies reflected a balanced, reproducible, and transparent selection process.

4.2 Quality Appraisal

We included 31 studies in the review, with methodological quality ranging from high to low. Each study was appraised against predefined quality criteria and rated as high, moderate, or low. Of these, 15 (48.4%) were rated as high quality, 10 (32.3%) as moderate quality, and 6 (19.4%) as low quality. High-quality studies have rigorous research designs and execution, whereas low-quality studies have notable weaknesses. Moderate-quality studies met some, but not all, the quality criteria.

Half of the included studies were rated as high quality, indicating robust evidence. These high-quality studies (Chan et al., 2024) employed strong experimental or quasi-experimental designs with rigorous control. They had random or well-matched participant assignments, appropriate comparison groups, and validated instruments to measure the learning outcomes. These studies reported more reliable and generalisable results than previous studies. In contrast, six studies were deemed low quality (Chun et al., 2025; Okubo et al., 2023) owing to their methodological weaknesses. The common weaknesses are as follows:

- Small samples or narrow contexts: Many studies had limited sample sizes or were conducted in a single class or institution, undermining their statistical power and external validity.
- Design or execution flaws: Several studies lacked randomisation or adequate control groups, increasing the risk of selection bias and confounding variables. Some participants or assessors were not blinded to the conditions, which could have introduced measurement bias.
- Limited outcome measures: Few low-quality studies relied on self-reported outcomes or unvalidated measures, thus providing weaker evidence of learning gain.

The remaining studies were of moderate quality (El-Sabagh, 2021; Abdulla et al., 2024), meeting some methodological criteria but falling short of others. These studies featured reasonable designs but had limitations, such as non-random assignments, moderate attrition, or less rigorous outcome measurements.

Higher methodological quality generally produces more credible findings. Our analysis showed that high-quality studies ($n = 16$) were more likely to report positive outcomes with clearer evidence of AI benefits.

RQ1: What evidence exists regarding the use of generative AI and adaptive systems to customise help-seeking scaffolds for learners in educational settings?

The 31 included studies spanned K-12, higher education, and adult learning contexts. Study settings ranged from controlled laboratory experiments to naturalistic classroom

deployments, with diverse learner populations.

Studies employed various AI/adaptive approaches: adaptive learning systems, conversational agents, intelligent tutoring systems, and LLM-based tools. Approaches reflected different disciplinary traditions and learner contexts.

Many studies used adaptive learning systems with analytics-driven personalisation, while others implemented conversational AI and dialogue systems. These diverse approaches reflected different disciplinary traditions and learner contexts.

Building on this typology, the subsequent analysis examined how each category of AI-driven scaffolds contributes to distinct phases of Self-Regulated Learning (SRL) and influences measurable learning outcomes (Li et al. (2025)). Based on the systematic SRL coding of all 31 studies (see Section 3.7 and Appendix A), Table 4 presents the cross-tabulation of scaffold categories by SRL phase. The performance phase was the most frequently targeted, with all 31 studies featuring scaffolds that provided real-time feedback, error correction, or strategy adjustment during task execution. The forethought phase was targeted by six studies, predominantly those employing real-time analytics prompts and rule-based scaffolding ($k = 4$; Lim et al., 2023; Pape & Thomas, 2025), which prompted goal-setting and strategic planning before task engagement. The self-reflection phase was coded in seven studies, most frequently among NLP/LLM-based feedback tools ($k = 4$; Chan et al., 2024; Menekse et al., 2025) and learning analytics dashboards ($k = 2$; Okubo et al., 2023; Wang et al., 2023), which prompted post-task evaluation and strategy revision. This pattern reveals a notable concentration on performance-phase support across all scaffold categories, with comparatively fewer interventions explicitly targeting forethought or self-reflection. The finding suggests that current AI-driven scaffolds are primarily reactive rather than proactive, responding to learner difficulties during task execution rather than promoting the planning and reflective processes that characterise self-regulated learners (Zimmerman, 2002). By grounding this typology in transparent coding rather than post hoc assertion, this review advances an integrated understanding of how scaffold design mediates the relationship between AI adaptability, learner agency, and academic performance (Panwale and Vijayakumar 2025; Wu et al., 2017).

Table 3

Summarises the major categories of AI-driven scaffold customisations observed (RQ1), along with example research studies.

AI-Driven Scaffold	Description	Citation of Studies
Real-Time Analytics Prompts / Rule-Based Scaffolding	Automated prompts triggered by log-data analysis of learner behavior (e.g., lack of planning) or fixed time intervals to support self-regulated learning. These use "if-then" rules to provide adaptive support that fades if actions are already taken.	Lim et al. (2023); Li et al. (2025); Lim et al. (2024); Pape & Thomas (2025)
NLP/LLM-Based Feedback, Personas & Chatbots	Usage of Natural Language Processing or Large Language Models (e.g., GPT-4, Llama 3.1) to generate feedback, summaries, or quality ratings based on student text/speech. This includes conversational agents using RAG (Retrieval-Augmented Generation) and systems assuming personas (e.g., Batman) to explain concepts or critique work.	Abdulla et al. (2024); Chan et al. (2024); Chun et al. (2025); Coban et al. (2025); Vicente-Yagüe-Jara et al. (2023); Dillon (2024); Menekse et al. (2025); Pesovski et al. (2024); Santhosh et al. (2024); Yan et al. (2025)
Adaptive Learning Systems & Games	Systems that adjust content difficulty, learning paths, or game mechanics based on real-time performance and "learning trajectories". Examples include expert systems using fuzzy logic and decision trees to personalize gaming pathways and environments that route learners based on VARK learning styles.	Bang et al. (2023); ElSabagh (2021); Horvers et al. (2024); Panjaburee et al. (2024); Panwale & Vijayakumar (2025); Reinhold et al. (2020)
Intelligent Tutoring Systems (ITS)	Advanced systems like TuinLECweb or SKOPE-IT that use virtual conversational agents to model strategies, provide explicit instruction, and offer adaptive elaborated feedback during problem-solving impasses.	Nye et al. (2018); Serrano-Mendizabal et al. (2023)
Learning Analytics Dashboards & Recommenders	Visual data displays offering descriptive (status updates) or prescriptive (suggested paths) feedback. Some function as review recommenders, suggesting specific pages based on quiz errors, peer signals, and "not understand" indicators.	Okubo et al. (2023); Wang et al. (2023)
AI-Generated Instructional Media & Tools	Includes AI-generated videos (created via platforms like Visla or Jasper) to teach scientific concepts and AI Writing Assistants (QuillBot) providing grammar, paraphrasing,	El-Garawany (2024); Pellas (2025)

AI-Driven Scaffold	Description	Citation of Studies
	and summarizing support for collaborative writing.	
Smart Personal Assistant (SPA)	Voice-activated AI (e.g., Amazon Alexa) providing dynamic scaffolding and problem-solving support in real-time.	Winkler et al. (2021)
Just-in-Time Discovery & Digital Scaffolding	Context-sensitive prompts provided during simulation tasks to guide inquiry strategies (predicting, testing) or digital environments providing on-demand cognitive and metacognitive hints.	Hajian et al. (2021); ter Beek et al. (2019); Wu et al. (2017)
Game-Based Incentive System	Automated mechanisms within educational games to reward engagement with content supports and regulate solution viewing.	Rahimi et al. (2021)

Table 4

Cross-tabulation of scaffold categories by SRL phase (Zimmerman, 2002). Each cell indicates the number of studies (k) in which the scaffold type was coded as targeting that phase. Studies could be coded to multiple phases.

Scaffold Category	Forethought (k)	Performance (k)	Self-Reflection (k)	Total Studies
Real-Time Analytics Prompts / Rule-Based Scaffolding	4	4	0	4
NLP/LLM-Based Feedback, Personas & Chatbots	1	10	4	10
Adaptive Learning Systems & Games	1	6	0	6
Intelligent Tutoring Systems (ITS)	0	2	0	2
Learning Analytics Dashboards & Recommenders	0	2	2	2
AI-Generated Instructional Media & Tools	0	2	1	2
Smart Personal Assistant (SPA)	0	1	0	1
Just-in-Time Discovery & Digital Scaffolding	0	3	0	3
Game-Based Incentive System	0	1	0	1
Total unique studies	6	31	7	31
<i>Note. k = number of studies. Column totals exceed 31 because studies coded to multiple phases are counted in each applicable column. Row totals reflect the number of unique studies per scaffold category.</i>				
RQ2: Compared with standard or non-customised scaffolds, how do AI-driven customised scaffolds impact learner engagement, academic performance, and satisfaction?				
Academic Performance				

AI-driven customised scaffolds generally had a positive impact on academic performance compared with standard scaffolding, though results were not universally positive. Of the 22 studies measuring learning outcomes, most showed improvements with AI-customised support. Notable findings include: greater quiz score improvements with adaptive review

dashboards (Okubo et al., 2023), improved learning gains with ChatGPT feedback (Coban et al., 2025), and better performance in personalised learning modules (Panwale & Vijayakumar, 2025). Lower-achieving students benefited most, with adaptive software showing substantial gains for these learners while high-achieving students showed mixed results.

Some studies reported no significant differences between AI-driven and control conditions. These null or mixed findings show the context-dependent nature of AI effectiveness and the importance of careful implementation.

In contrast, such support benefits those who need guidance, as seen in the moderation analyses by Pape and Thomas (2025). Most studies have indicated the positive effects of AI-customised scaffolds on academic performance. However, some studies have found neutral outcomes, demonstrating the importance of context and individual differences.

Learner Engagement

In six studies, AI scaffolding improved engagement-related outcomes including effort, motivation, and persistence. Students generally reported positive experiences with AI-driven support, though satisfaction varied by implementation quality. (Santhosh et al., 2024). These examples suggest that personalised scaffolding can encourage learners to interact more with the material and seek help when needed, thereby improving their engagement.

However, some studies have revealed several potential drawbacks to this approach. Pape and Thomas (2025) observed that students receiving adaptive prompts progressed more slowly through lessons than those without prompts, possibly indicating lower behavioural engagement or heavier cognitive load. Although disengagement was not directly measured, this suggests that poorly calibrated adaptive scaffolds may frustrate learners. Overall, the trend favours increased engagement with AI-tailored scaffolds, as learners respond positively to support tuned to their needs, leading to more time on tasks and deeper cognitive engagement than standard support.

Learner Satisfaction

Learner satisfaction and perceptions have been reported less frequently than performance and engagement outcomes. However, the results were generally positive for all the models. Few studies have explicitly evaluated students' satisfaction with their learning experiences. In an early implementation, Wu et al. (2017) found that participants in a dynamically scaffolded Excel training system expressed higher satisfaction and perceived adaptive scaffolds as helpful and motivating compared to traditional instruction. Similarly, Abdulla et al. (2024) reported that students using a ChatGPT-based coding assistant

reacted favourably, noting that on-demand AI support improved their confidence in problem-solving. These findings support the notion that personalised scaffolds can support a sense of support and competence, thereby improving learners' subjective satisfaction with the learning experience.

However, the evidence across studies remains inconsistent, and satisfaction does not always differ significantly from traditional approaches. For example, Chun et al. (2025) included "learning satisfaction" as an outcome measure and found no statistically significant difference between AI-driven and conventional textbook conditions, suggesting that factors such as usability, task relevance, and learner expectations may play a greater role than the presence of AI. Given these mixed findings, satisfaction currently provides only partial and indirect indications of professional relevance and the potential adoption of AI-driven scaffolds.

To strengthen this area, future research should develop more robust and standardised measures of learner satisfaction that go beyond general enjoyment or preferences and link satisfaction to perceived practical value, long-term engagement, and persistence in using AI support tools, which are critical for sustained adoption and educational impact. Researchers should also incorporate qualitative methods such as in-depth interviews, focus groups, or open-ended feedback mechanisms to explore *why* learners feel satisfied or dissatisfied with AI scaffolds. Such qualitative insights can reveal how factors such as perceived usefulness, trust, cognitive workload, and quality of AI feedback influence learner acceptance. A deeper understanding of these perceptions would clarify the mechanisms underlying satisfaction and inform the design of pedagogically effective, trustworthy, and user expectations, thereby improving their long-term relevance in educational contexts.

Table 5
Overview of reported impacts on engagement, performance and satisfaction.

Outcome	N of studies	General Impact of AI-Driven Scaffolds vs. Control	Citation of Studies
Academic Performance / Knowledge Gain / Achievement	27	Mixed. Positive effects were found in game-based, proactive agent, and dynamic scaffolding contexts. However, no significant main effect or direct difference was found in some rule-based analytics contexts, passive agent setups, or AI textbook comparisons.	Abdulla et al. (2024); Bang et al. (2023); Chan et al. (2024); Chun et al. (2025); Coban et al. (2025); El-Garawany (2024); Hajian et al. (2021); Horvers et al. (2024); Li et al. (2025); Lim et al. (2023); Lim et al. (2024); Nye et al. (2018); Okubo et al. (2023); Panjaburee et al. (2024); Panwale & Vijayakumar (2025); Pape & Thomas (2025); Pellas (2025); Pesovski et al. (2024); Rahimi et al. (2021); Reinhold et al. (2020); Santhosh et al. (2024); Serrano-Mendizabal et al. (2023); ter Beek et al. (2019); Wang et al. (2023); Winkler et al. (2021); Wu et al. (2017); Yan et al. (2025)
Self-Regulated Learning (SRL) / Metacognitive Activities	9	Mixed/Positive. Adaptive/Personalized scaffolds often induced higher frequencies of monitoring and regulatory activities. These process-level gains were more consistent than outcome gains, as they did not always translate to performance improvements.	Chun et al. (2025); Li et al. (2025); Lim et al. (2023); Lim et al. (2024); Pape & Thomas (2025); Rahimi et al. (2021); Serrano-Mendizabal et al. (2023); Wang et al. (2023); Winkler et al. (2021)
Engagement, Affect & Motivation	12	Positive. Adaptive systems and LLM feedback consistently improved motivation, engagement, self-efficacy, and confidence. Students reported higher satisfaction and reduced apprehension.	Bang et al. (2023); Chan et al. (2024); Dillon (2024); El-Garawany (2024); ElSabagh (2021); Panjaburee et al. (2024); Pellas (2025); Pesovski et al. (2024); Rahimi et al. (2021); Santhosh et al. (2024);

Outcome	N of studies	General Impact of AI-Driven Scaffolds vs. Control	Citation of Studies
			ter Beek et al. (2019); Wu et al. (2017)
Reflection & Writing Quality	2	Positive. AI feedback significantly improved the specificity of written reflections and increased essay revision scores.	Chan et al. (2024); Menekse et al. (2025)
Satisfaction / Usability Perceptions	3	Generally neutral-to-positive. Usability was rated strong in AR contexts, and satisfaction levels for AI textbooks were high and similar to traditional formats.	Chun et al. (2025); Coban et al. (2025); Okubo et al. (2023)
Learning Process Measures (Gaze / Logs)	2	Mixed. AI can reallocate attention toward task-relevant areas (eye-tracking), though log-based behavior results were not uniformly improved by prompting.	Coban et al. (2025); Pape & Thomas (2025)
Creativity (Fluency/Flexibility)	1	Positive (Hybrid). AI systems outperformed humans in fluency; the combination of humans and AI performed better than humans alone.	Vicente-Yagüe-Jara et al. (2023)
Goal Setting & Regulation of Practice	1	Positive. AI co-regulation improved regulation of practice (accuracy and problem volume) and goal-setting calibration.	Horvers et al. (2024)

RQ3: What are the standard methodologies and study designs used to evaluate the effectiveness of generative AI and adaptive systems in customising scaffolds?

Study Design Approaches: The included studies primarily employed experimental and quasi-experimental designs to assess the effectiveness of AI-driven scaffold customisation, utilising diverse, validated measurement instruments to evaluate learning outcomes. Most investigations employed a comparative group design, with an experimental group receiving AI-customised scaffolds compared to a control group receiving standard or no scaffolding. Approximately one-third of the studies were formal randomised controlled trials (RCTs) with randomly assigned participants. Chun et al. (2025) used a randomised two-group design (AI-improved vs. traditional textbooks) to compare learning outcomes in health science education. Pape and Thomas (2025) conducted an RCT testing adaptive

metacognitive prompting, measuring self-regulation using the Motivated Strategies for Learning Questionnaire (MSLQ) metacognitive self-regulation scale (12 items, Cronbach's $\alpha = 0.74$) and Time Management scale (eight items, $\alpha = 0.69$), along with behavioural trace data via LogRocket (capturing login duration, frequency, learning progress, and regularity) and custom recall/transfer tests with high inter-rater reliability ($\kappa = 0.99$).

Most studies ($n = 18$) used quasi-experimental designs with non-random assignment. Few used randomised controlled trials ($n = 4$), while qualitative and mixed methods studies ($n = 9$) provided contextual insights.

A few studies have involved single-group interventions without control groups, relying only on pre- and post-test comparisons. For example, Akbulut et al. (2024) evaluated learning gains before and after the use of AI-generated instructional videos, using domain-specific achievement measures. Although these studies have shown learning improvements, their designs limit their ability to make such causal inferences. Notably, validated cognitive load instruments (for example, the Paas Mental Effort Scale) and comprehensive self-efficacy scales were largely absent from these studies, representing a methodological gap in assessing the full theoretical mechanisms through which AI scaffolds operate in learning environments.

However, there were design variations between the studies. Most studies have employed between-subjects comparisons; however, Coban et al. (2025) used a within-subject crossover design in which groups alternated between using and not using AI feedback, allowing the participants to experience both conditions. Some studies have compared multiple treatment conditions rather than treatment versus control. Li et al. (2025) used three groups (no scaffolding, fixed static scaffolding, and adaptive scaffolding) to evaluate the benefits of adaptation. Serrano-Mendizábal et al. (2023) compared two scaffold types (strategy-focused vs. practice-focused) to determine which improved reading comprehension. These multigroup designs provide insights into the effectiveness of scaffolding strategies.

Table 6
Summary of the distribution of study design types in the included papers.

Study Design	N of Studies	Citation of Studies
Experimental (including Lab, Field, Pre-Post, Randomized, and 3-Condition Comparison)	12	Chan et al. (2024); Chun et al. (2025); Hajian et al. (2021); Li et al. (2025); Lim et al. (2023); Lim et al. (2024); Okubo et al. (2023); Pape & Thomas (2025); Santhosh et al. (2024); Serrano-Mendizabal et al. (2023); Wang et al. (2023); Wu et al. (2017)
Quasi-Experimental (including Pretest-Posttest Control Group)	11	Abdulla et al. (2024); Vicente-Yagüe-Jara et al. (2023); El-Garawany (2024); ElSabagh (2021); Horvers et al. (2024); Menekse et al. (2025); Nye et al. (2018); Panjaburee et al. (2024); Panwale & Vijayakumar (2025); ter Beek et al. (2019); Winkler et al. (2021)
Randomized Controlled Trial (RCT) (including Cluster and Lab-based RCT)	3	Bang et al. (2023); Reinhold et al. (2020); Yan et al. (2025)
Exploratory / Survey-Based / Case Study	2	Dillon (2024); Pesovski et al. (2024)
One-Group Pretest-Posttest / Within-Subjects	2	Rahimi et al. (2021); Pellas (2025)
Crossover Design	1	Coban et al. (2025)

Data Collection and Measures: Studies employed knowledge/skill assessments, self-reported help-seeking, and academic performance measures. Assessment methods varied considerably across studies, limiting direct comparison.

Most studies used statistical comparisons (t-tests, ANOVAs) or qualitative analysis. Effect sizes varied, from small ($d < 0.3$) to medium-to-large ($d > 0.8$), with heterogeneous patterns across contexts.

RQ4: What gaps and limitations have been reported in studies on AI-driven help-seeking scaffold customisation in education?

Across the studies, the authors acknowledged several limitations and gaps, often related to the scope and validity of their results. These limitations, summarised thematically below, point to the need for caution in interpreting the results and suggest directions for future research.

Sample Size and Generalizability

The most cited limitation was narrow study scope. Approximately 18 studies mentioned scope constraints (specific domains, learner populations, or AI tools), 12 reported small sample sizes, and 8 cited insufficient follow-up measures.

Short Duration and Lack of Long-Term Follow-Up

Many interventions have been tested in the short term (ranging from a single session to one academic term), and studies have explicitly mentioned the absence of longitudinal data as a limitation. Only immediate post-test outcomes were measured in this study. Therefore, it remains unclear whether the benefits of AI-driven scaffolds persist or translate into long-term skill development. El-Sabagh (2021) noted that there is no long-term follow-up to determine whether engagement gains lead to sustained academic improvement. Similarly, Okubo et al. (2023) evaluated dashboards after one review session and found that a short timeframe did not capture long-term study habits or knowledge retention. Without extended observation, we lack evidence on whether personalised scaffolding has persistent effects on learners' help-seeking behaviour and performance (a gap that future research should address through longitudinal studies or delayed post-tests).

Measurement and Instrumentation Limitations

Issues with measurement methods across studies included inconsistent instrumentation, limited validation of SRL-related measures, and lack of standardised outcome metrics. These inconsistencies limit synthesis and generalisation.

Regarding the study design, the papers acknowledged the methodological limitations. For example, Serrano-Mendizábal et al. (2023) noted the lack of a true "no-scaffold" control group, meaning that they compared the two scaffold types without isolating the effects of adaptive support. Such design limitations affect the interpretation of results and are noted as areas for future research improvement (e.g. including proper controls or confirming pre-test group equivalence).

Technical and Implementation Challenges

Limitations related to the technology and implementation of AI scaffolding have also been reported. One key concern is the reliance on digital infrastructure, which exacerbates the digital divide. Chun et al. (2025) noted that students without stable Internet access or adequate devices may be excluded from AI-based interventions, reinforcing existing educational inequities. To address this, several studies and policy reports have suggested practical approaches, such as developing low-bandwidth or offline-compatible AI tools, embedding AI functionalities within existing institutional learning management systems (LMSs) to increase accessibility, and supporting institutional and governmental initiatives

that fund digital inclusion for disadvantaged learners. Integrating AI scaffolds into familiar educational platforms also mitigates usability issues by reducing cognitive and technical barriers to adoption.

Usability challenges in emerging technologies can hinder adoption. Few studies explicitly addressed interface design or user experience issues in AI-driven scaffolds.

Ethical and Privacy Considerations

Although not the focus of most evaluations, several studies have demonstrated ethical issues and unintended consequences related to privacy, bias, and academic integrity. AI-driven scaffolds often rely on extensive data collection, including learner interactions, performance metrics, and behavioural analytics, which raises critical concerns regarding data storage, consent, and usage. Chun et al. (2025) and Abdulla et al. (2024) showed that opaque data pipelines and model biases may disproportionately affect specific groups of learners. To safeguard privacy, future implementations should adopt data minimisation and anonymisation protocols, use secure data encryption standards, and ensure institutional compliance with ethical governance frameworks for educational AI.

Therefore, academic integrity is equally important. Abdulla et al. (2024) cautioned that students might overrely on AI tools to generate rather than construct knowledge. To address this, educators and developers should design AI scaffolds as guided learning aids, not as solution providers, by embedding pedagogical prompts that promote reflection, self-explanation, and justification of AI-generated responses. In addition, clear usage policies and academic honesty guidelines should accompany the integration of AI tools into curricula to support responsible and transparent use.

Taken together, these considerations show that although technical and ethical challenges pose barriers to the professional adoption of AI-driven scaffolds, they are not insurmountable. Building low-resource AI tools, validating algorithmic accuracy, embedding human oversight, and enforcing privacy and integrity standards are essential steps towards equitable, trustworthy, and sustainable AI integration in education. Addressing these factors systematically ensures that the benefits of AI scaffolding extend beyond technological innovation to meaningful, inclusive, and ethical educational results.

Table 7
Common limitations reported across the studies

Reported Limitation	Studies (count)	Citation of Studies Noting this limitation
Small Sample Size / Limited Power / Attrition	17	Abdulla et al. (2024); Coban et al. (2025); Dillon (2024); El-Garawany (2024); Hajian et al. (2021); Li et al. (2025); Lim et al. (2023); Okubo et al. (2023); Panjaburee et al. (2024); Panwale & Vijayakumar (2025); Pape & Thomas (2025); Pellas (2025); Pesovski et al. (2024); Santhosh et al. (2024); Serrano-Mendizabal et al. (2023); Winkler et al. (2021); Wu et al. (2017)
Short Duration / Single Session / Novelty Effect / Sustainability	16	Chun et al. (2025); Coban et al. (2025); Vicente-Yagüe-Jara et al. (2023); Dillon (2024); El-Garawany (2024); Horvers et al. (2024); Li et al. (2025); Lim et al. (2023); Okubo et al. (2023); Panjaburee et al. (2024); Pape & Thomas (2025); Pellas (2025); Reinhold et al. (2020); Santhosh et al. (2024); Winkler et al. (2021); Yan et al. (2025)
Context Specificity / Limited Generalizability / Sample Homogeneity	14	Abdulla et al. (2024); Chan et al. (2024); Chun et al. (2025); Coban et al. (2025); Vicente-Yagüe-Jara et al. (2023); ElSabagh (2021); Li et al. (2025); Lim et al. (2023); Okubo et al. (2023); Panjaburee et al. (2024); Panwale & Vijayakumar (2025); Pape & Thomas (2025); Santhosh et al. (2024); Wang et al. (2023)
Study Design Issues (Lack of Control, Contamination, or Selection Bias)	10	Bang et al. (2023); Chun et al. (2025); Coban et al. (2025); Dillon (2024); El-Garawany (2024); Nye et al. (2018); Okubo et al. (2023); Pape & Thomas (2025); Pellas (2025); Rahimi et al. (2021)
Measurement, Technical, or Usability Constraints	7	Li et al. (2025); Lim et al. (2024); Menekse et al. (2025); Okubo et al. (2023); Pesovski et al. (2024); Reinhold et al. (2020); Serrano-Mendizabal et al. (2023)
Intervention Scope / Design Assumptions / Lack of Qualitative Data	5	ElSabagh (2021); Lim et al. (2023); Okubo et al. (2023); Pape & Thomas (2025); Serrano-Mendizabal et al. (2023)
AI and Technological Risks (Output accuracy, Algorithmic bias, Model dependence)	2	Chun et al. (2025); Coban et al. (2025)
Participant Behavioral Issues (Low engagement/uptake)	2	Nye et al. (2018); ter Beek et al. (2019)
Ecological Validity / Device and Access	2	Chun et al. (2025); Yan et al. (2025)

Reported Limitation	Studies (count)	Citation of Studies Noting this limitation
Inequity		

In short, research on AI tools that assist students in seeking help is growing and has yielded positive outcomes. However, there are still some significant limitations to this study. Many studies are in their early stages, with small groups, short timeframes and mixed results. This finding shows the need for more detailed long-term research. To improve this, researchers should use larger groups, study different settings, extend study times, better measure complex skills, improve AI tools, and consider ethical issues. These gaps provide researchers with an opportunity to develop more effective ways to utilise AI to support students in seeking help and learning independently of their teachers and peers.

Discussion

5.1 Effectiveness of AI-Driven Scaffolds

Studies revealed both benefits and challenges. Promise lies in personalised, responsive support; challenges include maintaining accuracy, maintaining engagement, and equitable access for diverse learners. These findings suggest that AI-driven scaffolds, ranging from intelligent tutoring systems to adaptive dashboards, can improve students’ mastery of content and encourage productive engagement in learning tasks.

AI scaffolding effectiveness varies by context. Some studies found no significant advantages over traditional instruction, while others showed improvements in specific domains. Variables affecting outcomes include learner prior knowledge, domain complexity, and scaffold design. Heterogeneous results suggest that AI scaffolds are not uniformly effective but require careful design and context-specific implementation.

5.2 Implications of Generative AI Tools

A notable trend in the recent literature is the integration of generative AI technologies into educational scaffolding. These tools offer new capabilities for personalised support, while introducing new challenges. Generative AI enables more dynamic and conversational forms of help that are not possible with earlier static systems (Mittal et al., 2024). Studies have demonstrated their effectiveness. ChatGPT provides on-demand explanations in AR science laboratories (Coban et al., 2025) and generates feedback on students' essays. Chan et al. (2024) found that students who received AI-generated feedback produced revisions with improved writing qualities. Coban et al. (2025) demonstrated that an AI tutor (GPT-4) could adapt its feedback, resulting in learning gains for students. In a programming course, students given access to ChatGPT as a

supplemental tutor showed higher performance, with many valuing AI support (Abdulla et al., 2024). These findings demonstrate the potential of generative AI to provide timely assistance that improves performance and learner satisfaction.

Generative AI in help-seeking contexts has significant limitations, particularly around content reliability, user dependence, and equity. Key concerns include maintaining AI responses are accurate, preventing over-reliance, and maintaining accessibility.

In general, generative AI tools bring powerful new affordances to help-seeking scaffolds but must be implemented with attention to ethical and pedagogical considerations. When used thoughtfully, they can improve educational scaffolding; however, if used naively, they can introduce errors into the learning process and lead to misunderstandings. The evidence offers optimism regarding these tools while cautioning about their limitations.

5.3 Patterns in Instructional Design and Adaptivity

Beyond outcomes, studies revealed distinct patterns in AI implementation. Most focused on single-domain applications, with limited cross-domain or interdisciplinary work. Learner populations spanned K-12 to adult education, with varying emphasis on underrepresented groups.

Scaffolding content varies by domain and instructional context, from domain-specific explanations to general metacognitive prompts. This variation reflects different pedagogical needs and learner populations.

The instructional contexts for these scaffolds varied, although most were implemented as supplements to the existing online learning environments. Studies have embedded AI-driven scaffolds into course activities, such as learning management system (LMS) plugins, digital textbook features and assignment tools (Silva et al., 2024). This integration means that the scaffolds work in conjunction with teacher-led instruction (except in self-paced settings). Nearly all interventions were evaluated in controlled settings, primarily in higher education courses rather than in K-12 classrooms. Tasks range from writing essays to solving math problems, conducting virtual labs with adaptive logic that assesses learner performance, and providing tailored support for learning (Isaeva et al., 2025). Emerging multimodal data for adaptivity were used in a study by Santhosh et al. (2024), who used gaze tracking to gauge engagement and trigger AI-generated reviews. Although such sensors are not yet mainstream, they suggest future designs in which AI can adjust to behavioural and affective cues, alongside performance metrics.

The design patterns in AI-driven help-seeking scaffolds range from straightforward personalisation to sophisticated continuous adaptivity through analytically driven recommendation. This field explores personalisation through content adaptation, strategic prompting and feedback customisation. Successful implementation aligns scaffolds closely

with learners' immediate needs and context (Alexsius Pardosi et al., 2024). Generic advice systems are less effective, and timely support relevant to student challenges is more engaging and helpful. This requires effective learner modelling and instructional design, leveraging AI techniques from decision-making rules to machine learning. These patterns offer educators an approach for designing AI-supported learning environments.

5.4 Methodological Reflections

Methodologically, studies employed diverse designs and measures, ranging from RCTs to naturalistic observation. This heterogeneity reflects developing field practices and context-specific research needs.

Most interventions were tested in the short term, ranging from a single session to one semester. Long-term studies tracking students across multiple semesters and courses are lacking. This limits the evidence on whether AI scaffolding benefits persist – do students maintain better help-seeking strategies after using an AI tutor, or do gains diminish? Short evaluation periods often fail to capture long-term outcomes, such as knowledge retention and skill transfer beyond immediate post-tests (Thomas et al., 2024).

Outcome measures were heterogeneous, including knowledge assessments, skill measurements, help-seeking behaviour, and self-regulated learning indicators. This diversity limits direct comparison across studies.

Despite limitations, few studies conducted rigorous internal validation or attempted replication. Only one study explicitly examined scalability; scalability remains a critical gap for future research. technological capabilities without a strong pedagogical foundation. This shows the opportunity to better connect AI interventions with learning science principles.

The methodological field is still developing, with evidence primarily from small-scale studies showing possibilities rather than large-scale trials demonstrating practical scalability. Variability in the measures and contexts suggests that future syntheses will require greater standardisation of methods. These limitations inform our review and indicate the following steps to strengthen evidence in this domain.

5.5 Limitations

Our search focused on peer-reviewed studies published in academic journals and conferences, primarily in English. Research in other languages or grey literature may have been missed, potentially skewing the findings towards positive outcomes, as unsuccessful experiments are less likely to be published. Our inclusion criteria focused on AI-driven help-seeking scaffolds, excluding broader AI education research and general AI tutoring systems unless they specifically addressed adaptive scaffolding in their use. Although

focused, this approach may have omitted relevant studies that were not explicitly framed as "help-seeking scaffolds". Theoretical papers and meta-analyses were excluded because we focused on empirical studies.

The heterogeneity of the studies prevented the meta-analytic synthesis. The interventions varied in terms of design and outcome measures, making it impossible to compute the aggregate effect sizes. We relied on qualitative synthesis, which introduces subjectivity into the interpretation of non-comparable evidence. We organised the findings thematically and cited specific examples, although this approach did not match the precision of a meta-analysis. Our statements regarding study frequency reflect judgments rather than quantitative analyses.

Resource constraints limit review processes, such as full double-anonymised screening and multiple independent reviewers. While we assessed the study quality to inform our interpretations, we did not exclude studies based on their quality. Therefore, studies with methodological limitations were included, which may have influenced these findings. Future reviews should focus on studies that meet quality thresholds or incorporate risk-of-bias analyses.

The timing of this review is significant, given the rapid evolution of AI in education, particularly in the context of generative AI (2022–2023). Our search covered literature up to 2025; however, new studies continue to emerge. The field of AI-driven educational tools is expanding, and technology will continue to advance in the future. These findings should be revisited periodically, as our review faces constraints of coverage, heterogeneity, and developing evidence, showing the need for further research.

Recommendations

Based on the findings and limitations of this review, we propose recommendations for research and practices/policies for educators, technology developers, and policymakers.

6.1 Research Recommendations

Expand Long-Term and Large-Scale Studies

Studies on AI-driven scaffolds should be conducted over extended periods in diverse populations. Future research should assess whether these improvements are sustained using longitudinal studies. Larger sample sizes and multisite studies should be conducted to increase the generalisability of these results. As noted by Chun et al. (2025), rigorous trials are required to evaluate the long-term impact and identify the cumulative effects of AI on learning.

Develop Standardised Metrics and Evaluation Frameworks

Standardised methods are required to measure key outcomes. Researchers should operationalise constructs such as "help-seeking behaviour" and "engagement" using standard questionnaires or logging schemas. A shared framework based on self-regulated learning measures facilitates evidence aggregation. The authors should report detailed usage patterns of AI interventions, including the frequency of help requests and types of errors, to enable systematic comparisons and meta-analyses.

Explore Under-Researched Contexts and Populations

Research has primarily focused on higher education and laboratory settings, with limited exploration of K-12, adult learning, and informal environments. Future studies should examine diverse educational contexts, including elementary and middle schools, to understand how AI-tailored scaffolding benefits students of different age groups. Research across various demographic populations and international education systems, particularly in low-resource settings, will reveal equity and implementation challenges that are not evident in university settings.

Investigate Process and Mechanisms

Further research is required to determine the effects of AI scaffolding on learning. Studies should include process measures, such as think-aloud protocols and trace data analysis, to examine student interactions. Mixed-methods approaches that combine quantitative outcomes with qualitative insights can reveal whether students prefer AI to human assistance and whether scaffold features affect help-seeking behaviours. Understanding these mechanisms will guide the design of artificial intelligence (AI) scaffolds to complement human instruction.

Address Technical and Ethical Dimensions in Research Designs

Therefore, researchers should extend their focus to the technical robustness and ethical implications of AI scaffolding, the technical aspects involved in testing different AI models, and their sensitivity to feedback errors. Ethical considerations included examining student trust, AI transparency, and the impact of content bias on students. Research has identified bias and privacy concerns, and studying these issues through bias mitigation trials and comparisons of data use is essential. This approach ensures that the AI scaffolds are effective, safe, and acceptable to all users.

6.2 Practical and Policy Recommendations

Integrate Teacher Training and Support

Institutions should provide professional development for teachers on AI scaffolding tool operation and pedagogical alignment, including when to use AI versus personal intervention (Chun et al., 2025). Teacher communities of practice can support integration and help identify common pitfalls such as AI errors and student overreliance.

Establish Ethical Guidelines and Data Policies

Institutions should establish clear policies on data privacy, consent, and academic integrity for AI use in classrooms. Ethics review processes should evaluate tools for bias and educational alignment before adoption, with particular attention to equity across linguistically diverse student populations.

Invest in Infrastructure and Accessibility

Schools must assess infrastructure readiness (devices, connectivity, bandwidth) and ensure AI tools are accessible to students with disabilities through universal design principles. Backup plans for AI service failures are essential.

Curriculum and Content Alignment

AI scaffolding should be aligned with curriculum objectives, with teachers involved in configuring hints and feedback to complement existing instruction. Scaffolds were most effective when reinforcing rather than replacing course materials.

Continuous Monitoring and Iteration

Implementation should be iterative: pilot in one semester, collect usage data and student feedback, and scale up after resolving issues. Administrators should compare outcomes between classes with and without AI scaffolds to evaluate effectiveness.

Promote AI Literacy and Student Training

AI literacy should be incorporated into learning objectives so students learn to formulate effective queries, interpret AI feedback critically, and verify information. Educators can use AI scaffolds to teach when to seek help versus work independently, maintaining that AI leads to empowerment rather than dependency.

Expected Outcomes & Contributions

This systematic review examined AI-driven scaffold customisation in education settings. A survey of 31 studies revealed gaps in large-scale evidence and long-term follow-up, particularly in the K-12 setting. Our thematic synthesis of effectiveness, generative AI implications, design patterns, and methodology consolidates the findings into a knowledge

base, enabling researchers to identify parallels across domains such as STEM tutoring and writing feedback. This synthesis advances our understanding by exposing gaps and suggests that future work could frame studies on self-regulated learning models.

This review offers practical insights for educators implementing AI-customised scaffolding. Practitioners can apply the findings in implementations; for example, a university writing centre might pilot AI feedback tools based on the demonstrated success in various contexts. Understanding curriculum alignment helps to avoid the use of pedagogically unsuitable tools. Our analysis of generative AI provides an evidence-based perspective on its benefits (engagement and personalised support) and risks (inaccuracies and teacher oversight), thereby informing institutional guidelines. The recommendations emphasise the importance of teacher training and AI literacy for maintaining effective implementation.

This review provides an application framework based on organised themes that practitioners can use for AI-driven learning scaffolding. Key considerations include evidence of effectiveness, implications of generative AI, design strategy, and measures of success. This framework translates research into practical decisions, advances knowledge, and improves educational practice. By identifying successes and areas of exploration, this review establishes a foundation for innovation in AI-supported learning.

Declarations

Ethics Statement

This study did not include any raw data. This review was based on aggregated studies that received ethical approval. No additional ethical approval was obtained for this study.

Conflict of Interest

The results of this study do not have any conflicts of interest to declare.

Funding Sources

This study did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement

The systematically analysed papers are available from the corresponding author, upon request.

References

1. Abbas AMH, Ting C-Y, Ghauth KI (2022) User Experience Design Using Machine Learning: A Systematic Review. *IEEE Access* 10:51501–51514. <https://doi.org/10.1109/access.2022.3173289>
2. *Abdulla S, Ismail S, Fawzy Y, Elhaj A (2024) Using ChatGPT in teaching computer programming and studying its impact on students' performance. *Electron J e-Learning* 22(6):66–81. <https://doi.org/10.34190/ejel.22.6.3380>
3. Aditya G (2024) Understanding and Addressing AI Hallucinations in Healthcare and Life Sciences. *Int J Health Sci* 7(3):1–11. <https://doi.org/10.47941/ijhs.1862>
4. Akinyi GL, Oboko R, Muchemi L (2024) Learning analytics intervention using prompts and feedback for measurement of e-learners' socially-shared regulated learning. *Electron J e-Learning* 22(5):103–116. <https://doi.org/10.34190/ejel.22.5.3253>
5. Al-Ameri A, Al-Shammari W, Nappi M, Umer M, Pero C, Castiglione A (2024) Student Academic Success Prediction Using Learning Management Multimedia Data With Convolved Features and Ensemble Model. *J Data Inform Qual*. <https://doi.org/10.1145/3687268>
6. Alam TM, Stoica GA, Özgöbek Ö (2025) Asking the classroom with technology: a systematic literature review. *Smart Learn Environ* 12(1). <https://doi.org/10.1186/s40561-024-00348-z>
7. Alenezi A (2024) Evaluating the effectiveness of AI-powered adaptive learning systems in secondary schools. 4686–700. <https://doi.org/10.46328/ijonse.264>
8. Alexsius Pardosi VB, Nurdiana N, Sabur F, Xu S, Umurohmi U (2024) Implementation of an Artificial Intelligence Based Learning Management System for Adaptive Learning. *Al-Fikrah. Jurnal Manajemen Pendidikan* 12(1):149. <https://doi.org/10.31958/jaf.v12i1.12548>
9. Ali D, Nikfar M, Ali H, Boskabadi E, Ugwuoke J, Fatemi Y (2024) ChatGPT in Teaching and Learning: A Systematic Review. *Educ Sci* 14(6):643. <https://doi.org/10.3390/educsci14060643>
10. Andrade HL (2019) A critical review of research on student self-assessment. *Front Educ* 4:87. <https://doi.org/10.3389/feduc.2019.00087>
11. Anjum F, Raheem BR, Ghafar ZN (2025) The Impact of ChatGPT on improving Students' Motivation and Learning Engagement in Second Language Acquisition: Insights from Students. *J E-Learning Res* 3(2):1–11. <https://doi.org/10.33422/jelr.v3i2.679>

12. Askill-Williams H, Skrzypiec G, Lawson MJ (2011) Scaffolding cognitive and metacognitive strategy instruction in regular class lessons. *Instr Sci* 40(2):413–443. <https://doi.org/10.1007/s11251-011-9182-5>
13. Ayedoun E, Hayashi Y, Seta K (2020) Toward personalized scaffolding and fading of motivational support in L2 learner–dialogue agent interactions: An exploratory study. *IEEE Trans Learn Technol* 13(3):604–616. <https://doi.org/10.1109/TLT.2020.2989776>
14. Ayeni O, Osawaru B, Hamad N, Chisom O, Adewusi O (2024) AI in education: A review of personalized learning and educational technology. *GSC Adv Res Reviews* 18(2):261–271. <https://doi.org/10.30574/gscarr.2024.18.2.0062>
15. Azarian M, Shiferaw AT, Yu H, Stevik TK (2023) Do We Perform Systematic Literature Review Right? A Scientific Mapping and Methodological Assessment. *Logistics* 7(4):89. <https://doi.org/10.3390/logistics7040089>
16. Azevedo R, Hadwin AF (2005) Scaffolding self-regulated learning and metacognition: Implications for the design of computer-based scaffolds. *Instr Sci* 33(5):367–379. <https://doi.org/10.1007/s11251-005-1272-9>
17. Azevedo R, Dever D, Wiedbusch M, Trevors G, Duffy M, Taub M, Cerezo R, Bouchet F, Cloude E, Wortha F, Harley J (2022) Lessons Learned and Future Directions of MetaTutor: Leveraging Multichannel Data to Scaffold Self-Regulated Learning With an Intelligent Tutoring System. *Front Psychol* 13(e60331). <https://doi.org/10.3389/fpsyg.2022.813632>
18. Azevedo R, Mudrick NV, Taub M, Bradbury AE (2019) Self-Regulation in Computer-Assisted Learning Systems. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education* (pp. 587–618). chapter, Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108235631.024>
19. Bandura A (1986) Social foundations of thought and action: A social cognitive theory. Prentice-Hall
20. *Bang HJ, Li L, Flynn K (2023) Efficacy of an adaptive game-based math learning app to support personalized learning and improve early elementary school students' learning. *Early Childhood Educ J* 51(4):717–732. <https://doi.org/10.1007/s10643-022-01332-3>
21. Banihashem SK, Noroozi O, Khosravi H, Bergdahl N, Bond M (2025) A systematic mapping review at the intersection of artificial intelligence and self-regulated learning. *Int J Educational Technol High Educ* 22(1). <https://doi.org/10.1186/s41239-025-00548-8>
22. Benton L, Mavrikis M, Vasalou A, Joye N, Sumner E, Herbert E, Revesz A, Symvonis A, Raftopoulou C (2021) Designing for challenge in a large-scale adaptive literacy game for primary school children. *Br J Edu Technol* 52(5):1862–1880. <https://doi.org/10.1111/bjet.13146>

23. Bertizzolo L, Bossuyt P, Atal I, Ravaud P, Dechartres A (2019) Disagreements in risk of bias assessment for randomised controlled trials included in more than one Cochrane systematic reviews: a research on research study using cross-sectional design. *BMJ Open* 9(4):e028382. <https://doi.org/10.1136/bmjopen-2018-028382>
24. Bettayeb AM, Altayasinah S, Abu Talib AZ, M., Dakalbab F (2024) Exploring the impact of ChatGPT: conversational AI in education. *Front Educ* 9. <https://doi.org/10.3389/educ.2024.1379796>
25. Bit D, Biswas S, Nag M (2024) The Impact of Artificial Intelligence in Educational System. *Int Journal Sci Res Sci Technology* 11(4):419-427. <https://doi.org/10.32628/ijrst2411424>
26. Borg M et al (2021) Exploring the assessment list for trustworthy AI in the context of advanced driver-assistance systems. *Proceedings of the 2021 IEEE/ACM 2nd International Workshop on Ethics in Software Engineering Research and Practice (SEthics)* (pp. 5-12). IEEE. <https://doi.org/10.1109/SEthics52569.2021.00009>
27. Braithwaite J, Herkes J, Churruca K, Long JC, Pomare C, Boyling C, Bierbaum M, Clay-Williams R, Rapport F, Shih P, Hogden A, Ellis LA, Ludlow K, Austin E, Seah R, Mcpherson E, Hibbert PD, Westbrook J (2019) Comprehensive Researcher Achievement Model (CRAM): a framework for measuring researcher achievement, impact and influence derived from a systematic literature review of metrics and models. *BMJ Open* 9(3):e025320. <https://doi.org/10.1136/bmjopen-2018-025320>
28. Bucchiarone A, Bonetti F, Yigitbas E (2025) Leveraging self-adaptive systems and generative AI for personalizing educational serious games: Architecture and future challenges. *Proceedings of the 2025 IEEE/ACM 20th Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)* (pp. 103-110). IEEE. <https://doi.org/10.1109/SEAMS66627.2025.00019>
29. Byrt T, Bishop J, Carlin JB (1993) Bias, prevalence and kappa. *J Clin Epidemiol* 46(5):423-429. [https://doi.org/10.1016/0895-4356\(93\)90018-V](https://doi.org/10.1016/0895-4356(93)90018-V)
30. Cicchetti DV, Feinstein AR (1990) High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 43(6):551-558. [https://doi.org/10.1016/0895-4356\(90\)90159-M](https://doi.org/10.1016/0895-4356(90)90159-M)
31. Calvi L, Cristea AI (2002) Towards Generic Adaptive Systems: Analysis of a Case Study. *Adaptive Hypermedia and Adaptive Web-Based Systems*. https://doi.org/10.1007/3-540-47952-X_10
32. Câmara A, Roy N, Maxwell D, Hauff C (2021) Searching to learn with instructional scaffolding. *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval (CHIIR '21)* (pp. 209-218). Association for Computing Machinery. <https://doi.org/10.1145/3406522.3446012>

33. Caratiquit KD, Caratiquit LJC (2023) ChatGPT as an academic support tool on the academic performance among students: The mediating role of learning motivation. *J Social Humanity Educ* 4(1):21–33. <https://doi.org/10.35912/jshe.v4i1.1558>
34. Carr A, Luckin R, Avramides K, Yuill N (2011) Scaffolding Metacognitive Processes in the Ecolab: Help-Seeking and Achievement Goal Orientation. *Artificial Intelligence in Education*. https://doi.org/10.1007/978-3-642-21869-9_60
35. Cervin-Ellqvist M, Larsson D, Adawi T, Stöhr C, Negretti R (2021) Metacognitive illusion or self-regulated learning? Assessing engineering students' learning strategies against the backdrop of recent advances in cognitive science. *High Educ* 82(3):477–498. <https://doi.org/10.1007/s10734-020-00635-x>
36. *Chan S, Lo N, Wong A (2024) Generative AI and Essay Writing: Impacts of Automated Feedback on Revision Performance and Engagement. *rEFLections* 31(3):1249–1284. <https://doi.org/10.61508/refl.v31i3.277514>
37. Chang DH, Lin MP-C, Wang QQ, Hajian S (2023) Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization. *Sustainability* 15(17):12921. <https://doi.org/10.3390/su151712921>
38. Chen J, Chen G, Fu H, Lin C-H (2023) Individual differences in self-regulated learning profiles of Chinese EFL readers: A sequential explanatory mixed-methods study. *Stud Second Lang Acquisition* 45(4):955–978. <https://doi.org/10.1017/s0272263122000584>
39. Chen Y-C, Wu C-H, Hou H-T (2022) Design and Development of a Scaffolding-Based Mindtool for Gamified Learning Classrooms. *J Educational Comput Res* 61(1):3–29. <https://doi.org/10.1177/07356331221101081>
40. Chen Y, Lu J, Wang L, Gao X (2023) Effective interventions for gaming disorder: A systematic review of randomized control trials. *Frontiers in Psychiatry*, 14. <https://doi.org/10.3389/fpsy.2023.1098922>
41. Cheng HNH, Yang EFY, Liao CCY, Chang B, Huang YCY, Chan T-W (2015) Scaffold Seeking: A Reverse Design of Scaffolding in Computer-Supported Word Problem Solving. *J Educational Comput Res* 53(3):409–435. <https://doi.org/10.1177/0735633115601598>
42. Cheong I, Caliskan A, Kohno T (2024) Safeguarding human values: rethinking US law for generative AI's societal impacts. *AI Ethics* 5(2):1433–1459. <https://doi.org/10.1007/s43681-024-00451-4>
43. *Chun J, Kim J, Kim H, Lee G, Cho S, Kim C, Chung Y, Heo S (2025) A comparative analysis of on-device AI-driven self-regulated learning and traditional pedagogy in university health sciences education. *Appl Sci* 15(4) Article 1815. <https://doi.org/10.3390/app15041815>

44. *Coban A, Dzsoťjan D, Kuechemann S, Durst J, Kuhn J, Hoyer C (2025) AI support meets AR visualization for Alice and Bob: Personalized learning based on individual ChatGPT feedback in an AR quantum cryptography experiment for physics lab courses. *EPJ Quantum Technol* 12:15. <https://doi.org/10.1140/epjqt/s40507-025-00310-z>
45. Córdova-Esparza D-M (2025) AI-Powered Educational Agents: Opportunities, Innovations, and Ethical Challenges. *Information* 16(6):469. <https://doi.org/10.3390/info16060469>
46. Dagunduro A, Ajuwon O, Chikwe C, Ediae A (2024) Adaptive Learning Models for Diverse Classrooms: improving Educational Equity. *Int J Appl Res Social Sci* 6(9):2228–2240. <https://doi.org/10.51594/ijarss.v6i9.1588>
47. *Dillon T (2024) Korean university students' prompt literacy training with ChatGPT: Investigating language learning strategies. *Engl Teach* 79(3):123–157. <https://doi.org/10.15858/engtea.79.3.202409.123>
48. Doebling A, Kazerouni AM (2021) Patterns of academic help-seeking in undergraduate computing students. *Proceedings of the 21st Koli Calling International Conference on Computing Education Research* (Article 13, 10 pp.). Association for Computing Machinery. <https://doi.org/10.1145/3488042.3488052>
49. Ejjami R (2024) The Adaptive Personalization Theory of Learning: Revolutionizing Education with AI. *Journal of Next-Generation Research* 5.0. <https://doi.org/10.70792/jngr5.0.v1i1.8>
50. *El-Garawany MSM (2024) The effects of a QuillBot-based intervention on English language majors' EFL writing performance, apprehension, and self-efficacy. *Lang Teach Res Q* 43:167–189. <https://doi.org/10.32038/ltrq.2024.43.10>
51. *El-Sabagh HA (2021) Adaptive e-learning environment based on learning styles and its impact on developing students' engagement. *Int J Educational Technol High Educ* 18 Article 53. <https://doi.org/10.1186/s41239-021-00289-4>
52. Elsayad G (2024) Drivers of undergraduate students' learning perceptions in the blended learning environment: The mediation role of metacognitive self-regulation. *Educ Inform Technol* 29(12). <https://doi.org/10.1007/s10639-024-12466-9>
53. Faber TJE, Dankbaar MEW, Van Den Broek WW, Bruinink LJ, Hogeveen M, Van Merriënboer JJG (2024) Effects of adaptive scaffolding on performance, cognitive load and engagement in game-based learning: a randomized controlled trial. *BMC Med Educ* 24(1). <https://doi.org/10.1186/s12909-024-05698-3>
54. Faza A, Lestari IA (2025) Self-Regulated Learning in the Digital Age: A Systematic Review of Strategies, Technologies, Benefits, and Challenges. *Int Rev Res Open Distrib Learn* 26(2):23–58. <https://doi.org/10.19173/irrodl.v26i2.8119>
55. Ginting D, Barella Y, Woods R, Madkur A, Sabudu D, Sari MK (2024) Student-centered learning in the digital age: in-class adaptive instruction and best practices.

- International Journal of Evaluation and Research in Education (IJERE)*, 13(3), 2006.
<https://doi.org/10.11591/ijere.v13i3.27497>
56. Gomes D (2025) A Comprehensive Study of Advancements in Intelligent Tutoring Systems Through Artificial Intelligent Education Platforms. In F. Moreira & R. Teles (Eds.), *Improving Student Assessment With Emerging AI Tools* (pp. 213–244). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-6170-2.ch008>
 57. Gomez C, Huang C-M, Cho SM, Ke S, Unberath M (2025) Human-AI collaboration is not very collaborative yet: a taxonomy of interaction patterns in AI-assisted decision making from a systematic review. *Front Comput Sci* 6.
<https://doi.org/10.3389/fcomp.2024.1521066>
 58. Guettala M, Bourekkache S, Kazar O, Harous S (2024) Generative Artificial Intelligence in Education: Advancing Adaptive and Personalized Learning. *Acta Informatica Pragensia* 13(3):460–489. <https://doi.org/10.18267/j.aip.235>
 59. Guo L (2022) Using metacognitive prompts to improve self-regulated learning and learning outcomes: A meta-analysis of experimental studies in computer-based learning environments. *J Comput Assist Learn* 38(3):811–832.
<https://doi.org/10.1111/jcal.12650>
 60. Guo Y, Wang Y, Ortega-Martín JL (2023) The impact of blended learning-based scaffolding techniques on learners' self-efficacy and willingness to communicate. *Porta Linguarum Revista Interuniversitaria de Didáctica de Las Lenguas Extranjeras* 40:253–273. <https://doi.org/10.30827/portalin.vi40.27061>
 61. Guyatt G, Schandelmaier S, Brignardello-Petersen R, De Beer H, Prasad M, Murad MH, Eachempati P, Chu DK, D'Souza R, Iorio A, Agoritsas T, Yao L, Mustafa RA, Parpia S, Santaguida P, Vandvik PO, Hultcrantz M, Montori VM (2025) Core GRADE 3: rating certainty of evidence-assessing inconsistency. *BMJ (Clinical Res Ed)* 389:e081905. <https://doi.org/10.1136/bmj-2024-081905>
 62. *Hajian S, Jain M, Liu AL, Obaid T, Fukuda M, Winne PH, Nesbit JC (2021) improving scientific discovery learning by just-in-time prompts in a simulation-assisted inquiry environment. *Eur J Educational Res* 10(2):989–1007. <https://doi.org/10.12973/eu-er.10.2.989>
 63. Haroud S, Saqri N (2025) Generative AI in higher education: Teachers' and students' perspectives on support, replacement and digital literacy. *Educ Sci* 15(4). Article 396. <https://doi.org/10.3390/educsci15040396>
 64. He J, Wang J, Zhong H, Guan C (2022) The Effectiveness of Multi-Component Interventions on the Positive and Negative Aspects of Well-Being among Informal Caregivers of People with Dementia: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health* 19(12):6973. <https://doi.org/10.3390/ijerph19126973>

65. Hernandez-Cruz N, Sarker MMK, Saha P, Noble JA (2024) Review of Federated Learning and Machine Learning-Based Methods for Medical Image Analysis. *Big Data Cogn Comput* 8(9):99. <https://doi.org/10.3390/bdcc8090099>
66. Hill JE, Harris C, Clegg A (2023) Methods for using Bing's AI-powered search engine for data extraction for a systematic review. *Res Synthesis Methods* 15(2):347–353. <https://doi.org/10.1002/jrsm.1689>
67. Hirt J, Meichlinger J, Schumacher P, Mueller G (2021) Agreement in Risk of Bias Assessment Between RobotReviewer and Human Reviewers: An Evaluation Study on Randomised Controlled Trials in Nursing-Related Cochrane Reviews. *J Nurs Scholarsh* 53(2):246–254. <https://doi.org/10.1111/jnu.12628>
68. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, Pluye P (2022) The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inform* 38(2):69–82. <https://doi.org/10.3233/EFI-180221>
69. Hong S, Hwang J, Eom T, Moon J, Awoyemi ID (2025) Generative AI-improved Virtual Reality Simulation for Pre-Service Teacher Education: A Mixed-Methods Analysis of Usability and Instructional Utility for Course Integration. *Educ Sci* 15(8):997. <https://doi.org/10.3390/educsci15080997>
70. *Horvers A, Kooi R, Knoop-van Campen CAN, Dijkstra R, Baars M, Molenaar I (2024) How does co-regulation with Adaptive Learning Technologies affect primary school students' goal-setting, regulation of practice behavior and learning outcomes? *Front Educ* 9:1435483. <https://doi.org/10.3389/educ.2024.1435483>
71. Hou H-T, Keng S-H (2020) A Dual-Scaffolding Framework Integrating Peer-Scaffolding and Cognitive-Scaffolding for an Augmented Reality-Based Educational Board Game: An Analysis of Learners' Collective Flow State and Collaborative Learning Behavioral Patterns. *J Educational Comput Res* 59(3):547–573. <https://doi.org/10.1177/0735633120969409>
72. Hu K (2023), February 2 *ChatGPT sets record for fastest-growing user base - analyst note*. Reuters. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
73. Huang Z, Mao Y, Zhang J (2024) The Influence of Artificial Intelligence Technology on College Students' Learning Effectiveness from the Perspective of Constructivism — Taking ChatGPT as an Example. *J Educ Humanit Social Sci* 30:40–46. <https://doi.org/10.54097/y1x3jj43>
74. Ifraheem S, Siddiqui A, Rasheed M (2024) Transforming Education Through Artificial Intelligence: Personalization, Engagement and Predictive Analytics. *J Asian Dev Stud* 13(2):250–266. <https://doi.org/10.62345/jads.2024.13.2.22>
75. Isaeva R, Dznunusnalieva K, Mokliuk M, Karasartova N, Mirzoeva K (2025) improving learning effectiveness through adaptive learning platforms and emerging computer

- technologies in education. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi* 9(1):144-160. <https://doi.org/10.22437/jiituj.v9i1.37967>
76. Janssen M (2025) Responsible governance of generative AI: Conceptualizing GenAI as complex adaptive systems. *Policy Soc* 44(1):38-51. <https://doi.org/10.1093/polsoc/puae040>
 77. Järvelä S, Nguyen A, Molenaar I (2023) Advancing SRL research with artificial intelligence. *Comput Hum Behav* 147:107847. <https://doi.org/10.1016/j.chb.2023.107847>
 78. Jin FJ-Y, Maheshi B, Lai W, Li Y, Gašević D, Chen G, Charwat N, Chan PWK, Martinez-Maldonado R, Tsai Y-S (2025) Students' perceptions of generative AI-powered learning analytics in the feedback process: A feedback literacy perspective. *J Learn Analytics* 12(1):152-168. <https://doi.org/10.18608/jla.2025.8609>
 79. Jin F, Maheshi B, Martinez-Maldonado R, Gašević D, Tsai Y-S (2024) Scaffolding feedback literacy: Designing a feedback analytics tool with students. *J Learn Analytics* 11(2):123-137. <https://doi.org/10.18608/jla.2024.8339>
 80. Jin S-H, Yoo M, Roll I, Seo K, Im K (2023) Supporting students' self-regulated learning in online learning using artificial intelligence applications. *Int J Educational Technol High Educ* 20(1). <https://doi.org/10.1186/s41239-023-00406-5>
 81. Jumaat NF, Tasir Z (2014) Instructional scaffolding in online learning environment: A meta-analysis. In *Proceedings of the 2014 International Conference on Teaching and Learning in Computing and Engineering (LaTiCE)* (pp. 74-77). IEEE. <https://doi.org/10.1109/LaTiCE.2014.22>
 82. Karabenick SA, Newman RS (2006) *Help Seeking in Academic Settings: Goals, Groups, and Contexts* (1st ed.). Routledge. <https://doi.org/10.4324/9780203726563>
 83. Karlen Y, Hirt CN, Liska A, Stebner F (2021) Mindsets and Self-Concepts About Self-Regulated Learning: Their Relationships with Emotions, Strategy Knowledge, and Academic Achievement. *Front Psychol* 12. <https://doi.org/10.3389/fpsyg.2021.661142>
 84. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, Gasser U, Groh G, Günnemann S, Hüllermeier E, Krusche S, Kutyniok G, Michaeli T, Nerdel C, Pfeffer J, Poquet O, Sailer M, Schmidt A, Seidel T, Kasneci G (2023) ChatGPT for good? On opportunities and challenges of large language models for education. *Learn Individual Differences* 103:102274. <https://doi.org/10.1016/j.lindif.2023.102274>
 85. Khalil M, Paas F, Wasson B, Wong J (2024) Adaptive support for self-regulated learning in digital learning environments. *Br J Edu Technol* 55(4):1281-1289. <https://doi.org/10.1111/bjet.13479>
 86. Kochmar E, Gupta V, Serban IV, Pineau J, Belfer R, Vu DD (2021) Automated Data-Driven Generation of Personalized Pedagogical Interventions in Intelligent Tutoring

- Systems. *Int J Artif Intell Educ* 32(2):323–349. <https://doi.org/10.1007/s40593-021-00267-x>
87. Kong S-C, Yang Y (2024) A Human-Centered Learning and Teaching Framework Using Generative Artificial Intelligence for Self-Regulated Learning Development Through Domain Knowledge Learning in K-12 Settings. *IEEE Trans Learn Technol* 17:1562–1573. <https://doi.org/10.1109/tlt.2024.3392830>
88. Kusmaryono I, Jupriyanto, Kusumaningsih W (2021) Construction of students' mathematical knowledge in the zone of proximal development and zone of potential construction. *Eur J Educational Res* 10(1):341–351. <https://doi.org/10.12973/eu-jer.10.1.341>
89. Lai JW (2024) Adapting Self-Regulated Learning in an Age of Generative Artificial Intelligence Chatbots. *Future Internet* 16(6):218. <https://doi.org/10.3390/fi16060218>
90. Lan M, Zhou X (2025) A qualitative systematic review on AI empowered self-regulated learning in higher education. *Npj Sci Learn* 10(1). <https://doi.org/10.1038/s41539-025-00319-0>
91. Lee Y, Chung JJY, Kim TS, Song JY, Kim J (2022) Promptiverse: Scalable generation of scaffolding prompts through human-AI hybrid knowledge graph annotation. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (Article 96, 18 pp.). Association for Computing Machinery. <https://doi.org/10.1145/3491102.3502087>
92. Li M, Gao Q, Yu T (2023) Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters. *BMC Cancer* 23(1). <https://doi.org/10.1186/s12885-023-11325-z>
93. Li T, Raković M, Lim L, Li X, Yang B, Bannert M, Moore J, Tan Y, Tsai Y-S, Swiecki Z, Shaffer DW, Singh S, Wang Y, Van Der Graaf J, Fan Y, Gašević D, Molenaar I (2023) Analytics of self-regulated learning scaffolding: effects on learning processes. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1206696>
94. *Li T, Yan L, Iqbal S, Srivastava N, Singh S, Raković M, Swiecki Z, Tsai Y-S, Gašević D, Fan Y, Li X (2025) Analytics of self-regulated learning strategies and scaffolding: Associations with learning performance. *Computers Education: Artif Intell* 8:100410. <https://doi.org/10.1016/j.caeai.2025.100410>
95. Lim L-A, Gašević D, Matcha W, Ahmad Uzir NA, Dawson S (2021) Impact of learning analytics feedback on self-regulated learning: Triangulating behavioural logs with students' recall. *Proceedings of the 11th International Learning Analytics and Knowledge Conference (LAK '21)* (pp. 364–374). Association for Computing Machinery. <https://doi.org/10.1145/3448139.3448174>
96. *Lim L, Bannert M, van der Graaf J, Fan Y, Raković M, Singh S, Molenaar I, Gašević D (2024) How do students learn with real-time personalized scaffolds? *Br J Edu Technol*

- 55(4):1309–1327. <https://doi.org/10.1111/bjet.13414>
97. *Lim L, Bannert M, van der Graaf J, Singh S, Fan Y, Surendrannair S, Raković M, Molenaar I, Moore J, Gašević D (2023) Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Comput Hum Behav* 139:107547. <https://doi.org/10.1016/j.chb.2022.107547>
 98. Lim L, Tabatabai MM, Sufi S, Roy B, S., Loizou P (2023) Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Comput Hum Behav* 139:107547. <https://doi.org/10.1016/j.chb.2022.107547>
 99. Ling H-C, Chiang H-S (2022) Learning Performance in Adaptive Learning Systems: A Case Study of Web Programming Learning Recommendations. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.770637>
 00. Lipman JM, Colbert CY, Ashton R, French J, Warren C, Yepes-Rios M, King RS, Bierer SB, Kline T, Stoller JK (2023) A Systematic Review of Metrics Utilized in the Selection and Prediction of Future Performance of Residents in the United States. *J Graduate Med Educ* 15(6):652–668. <https://doi.org/10.4300/jgme-d-22-00955.1>
 01. Liu Z, Yin SX, Lee C, Chen NF (2024) Scaffolding Language Learning via Multi-modal Tutoring Systems with Pedagogical Instructions. 2024 IEEE Conference on Artificial Intelligence (CAI), Singapore, Singapore, 2024, pp. 1258–1265. <https://doi.org/10.1109/cai59869.2024.00223>
 02. Luo QZ (2023) The Influence of AI-Powered Adaptive Learning Platforms on Student Performance in Chinese Classrooms. *J Educ* 6(3):1–12. <https://doi.org/10.53819/81018102t4181>
 03. Mahan KR (2020) The comprehending teacher: scaffolding in content and language integrated learning (CLIL). *Lang Learn J* 50(1):74–88. <https://doi.org/10.1080/09571736.2019.1705879>
 04. Mahboob K, Asif R, Umme L (2024) Leveraging Generative AI for Cross-Cultural Knowledge Exchange in Higher Education. In P. Yu, J. Mulli, Z. Syed, & L. Umme (Eds.), *Facilitating Global Collaboration and Knowledge Sharing in Higher Education With Generative AI* (pp. 186–206). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-0487-7.ch008>
 05. Makridakis S, Petropoulos F, Kang Y (2023) Large Language Models: Their Success and Impact. *Forecasting* 5(3):536–549. <https://doi.org/10.3390/forecast5030030>
 06. Martin ND, Tissenbaum D, Gnesdilow C, D., Puntambekar S (2018) Fading distributed scaffolds: the importance of complementarity between teacher and material scaffolds. *Instr Sci* 47(1):69–98. <https://doi.org/10.1007/s11251-018-9474-0>
 07. Massaty MH, Budiyanto CW, Fahrurrozi SK (2024) The Role of AI in supporting Computational Thinking and Self-Efficacy in Educational Settings: A Systematic

- Review. *IJIE (Indonesian J Inf Education)* 8(1):49.
<https://doi.org/10.20961/ijie.v8i1.89596>
08. Mcgrath C, Farazouli A, Cerratto-Pargman T (2024) Generative AI chatbots in higher education: a review of an emerging research area. *High Educ* 89(6):1533–1549.
<https://doi.org/10.1007/s10734-024-01288-w>
 09. McHugh ML (2012) Interrater reliability: The kappa statistic. *Biochemia Med* 22(3):276–282. <https://doi.org/10.11613/BM.2012.031>
 10. Memon MQ, Shah SFA, Memon A, Lu Y, Memon AR, Munshi P (2022) Does the Impact of Technology Sustain Students’ Satisfaction, Academic and Functional Performance: An Analysis via Interactive and Self-Regulated Learning? *Sustainability* 14(12):7226.
<https://doi.org/10.3390/su14127226>
 11. *Menekse M, Putra AS, Kim J, Butt AA, McDaniel MA, Davidesco I, Cadieux M, Kim J, Litman D (2025) improving student reflections with natural language processing-based scaffolding: A quasi-experimental study in a large lecture course. *Computers Education: Artif Intell* 8:100397. <https://doi.org/10.1016/j.caeai.2025.100397>
 12. Miao J, Ma L (2023) Teacher Autonomy Support Influence on Online Learning Engagement: The Mediating Roles of Self-Efficacy and Self-Regulated Learning. *Sage Open* 13(4). <https://doi.org/10.1177/21582440231217737>
 13. Mittal U, Chamola V, Sai S, Sangwan D (2024) A Comprehensive Review on Generative AI for Education. *IEEE Access* 12:142733–142759.
<https://doi.org/10.1109/access.2024.3468368>
 14. Mohammed PS, ‘Nell’ Watson E (2019) Towards Inclusive Education in the Age of Artificial Intelligence: Perspectives, Challenges, and Opportunities. In: Knox, J., Wang, Y., Gallagher, M. (eds) *Artificial Intelligence and Inclusive Education. Perspectives on Rethinking and Reforming Education. Springer, Singapore*.
https://doi.org/10.1007/978-981-13-8161-4_2
 15. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009) Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 6(7):e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
 16. Moons F, Vandervieren E (2025) Measuring agreement among several raters classifying subjects into one or more (hierarchical) categories: A generalization of Fleiss’ kappa. *Behav Res Methods* 57(10). <https://doi.org/10.3758/s13428-025-02746-8>
 17. Morales-Chan M, Amado-Salvatierra HR, Medina JA, Barchino R, Hernandez-Rizzardini R, Teixeira AM (2024) Personalized feedback in massive open online courses: Harnessing the power of LangChain and OpenAI API. *Electronics*, 13(10), Article 1960.
<https://doi.org/10.3390/electronics13101960>

18. Munshi A, Baker R, Biswas G, Ocumpaugh J, Hutt S, Paquette L (2022) Analysing adaptive scaffolds that help students develop < scp>self-regulated learning behaviours. *J Comput Assist Learn* 39(2):351–368. <https://doi.org/10.1111/jcal.12761>
19. Nascimento LDS, Reichert FM, Janissek-Muniz R, Zawislak PA (2020) Dynamic interactions among knowledge management, strategic foresight and emerging technologies. *J Knowl Manage* 25(2):275–297. <https://doi.org/10.1108/jkm-01-2020-0044>
20. Naseer F, Khawaja S (2025) Mitigating Conceptual Learning Gaps in Mixed-Ability Classrooms: A Learning Analytics-Based Evaluation of AI-Driven Adaptive Feedback for Struggling Learners. *Appl Sci* 15(8):4473. <https://doi.org/10.3390/app15084473>
21. Navas Bonilla CDR, Noriega M, Pupiales DEG, J. C., Viñan Carrasco LM (2025) The Future of Education: A Systematic Literature Review of Self-Directed Learning with AI. *Future Internet* 17(8):366. <https://doi.org/10.3390/fi17080366>
22. Nikolovski V, Chorbev I, Trajanov D (2025) Advancing AI in Higher Education: A Comparative Study of Large Language Model-Based Agents for Exam Question Generation, Improvement, and Evaluation. *Algorithms* 18(3):144. <https://doi.org/10.3390/a18030144>
23. Nussbaumer-Streit B, Sommer I, Hamel C, Devane D, Noel-Storr A, Puljak L, Trivella M, Gartlehner G (2023) Rapid reviews methods series: Guidance on team considerations, study selection, data extraction and risk of bias assessment. *BMJ Evidence-Based Med* 28(6):418–423. <https://doi.org/10.1136/bmjebm-2022-112185>
24. *Nye BD, Pavlik PI Jr., Windsor AM, Olney AM, Fossati D (2018) SKOPE-IT (Shareable Knowledge Objects as Portable Intelligent Tutors): Overlaying natural language tutoring on an adaptive learning system for mathematics. *Int J STEM Educ* 5(1). Article 15. <https://doi.org/10.1186/s40594-018-0111-5>
25. *Okubo F, Shiino T, Minematsu T, Taniguchi Y, Shimada A (2023) Adaptive learning support system based on automatic recommendation of personalized review materials. *IEEE Trans Learn Technol* 16(1):92–105. <https://doi.org/10.1109/TLT.2022.3225206>
26. Ou S (2024) Transforming Education: The developing Role of Artificial Intelligence in The Students Academic Performance. *Int J Educ Humanit* 13(2):163–173. <https://doi.org/10.54097/cc1x7r95>
27. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A (2016) Rayyan—a web and mobile app for systematic reviews. *Syst Reviews* 5:210. <https://doi.org/10.1186/s13643-016-0384-4>
28. Paethrangsi N, Teekasap S, Jandaboue W, Khiewpan R (2024) Empowering Students' Autonomous Learning through Self-regulation, Metacognitive Strategies, and Collaborative Learning Environments. *J Liberal Arts RMUTT* 5(1):69–79. <https://doi.org/10.60101/jla.2024.5.1.4065>

29. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Moher D (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372:n71. <https://doi.org/10.1136/bmj.n71>
30. Paglialunga A, Melogno S (2025) The Effectiveness of Artificial Intelligence-Based Interventions for Students with Learning Disabilities: A Systematic Review. *Brain Sci* 15(8):806. <https://doi.org/10.3390/brainsci15080806>
31. Panadero E (2017) A review of self-regulated learning: Six models and four directions for research. *Front Psychol* 8:422. <https://doi.org/10.3389/fpsyg.2017.00422>
32. *Panjaburee P, Hwang G-J, Intarakamhang U, Srisawasdi N, Chaipidech P (2024) Effects of a personalized game on students' outcomes and visual attention during digital citizenship learning. *Cogent Educ* 11(1). Article 2351275. <https://doi.org/10.1080/2331186X.2024.2351275>
33. Panwale SB, Vijayakumar S (2025) Evaluating AI-personalized learning interventions in distance education. *Int Rev Res Open Distrib Learn* 26(1):158-174. <https://doi.org/10.19173/irrodl.v26i1.7813>
34. *Pape R, Thomas J (2025) Adaptive metacognitive prompting in young learners and the role of prior performance. *Int J Child-Comput Interact* 44 Article 100740. <https://doi.org/10.1016/j.ijcci.2025.100740>
35. *Pellas N (2025) The impact of AI-generated instructional videos on problem-based learning in science teacher education. *Educ Sci* 15(1) Article 102. <https://doi.org/10.3390/educsci15010102>
36. *Pesovski I, Santos R, Henriques R, Trajkovic V (2024) Generative AI for customizable learning experiences. *Sustainability* 16(7) Article 3034. <https://doi.org/10.3390/su16073034>
37. Pintrich PR (2000) The role of goal orientation in self-regulated learning. In: Boekaerts M, Pintrich PR, Zeidner M (eds) *Handbook of self-regulation: Research, theory, and applications*. Elsevier Science, pp 451-502
38. Prasad P, Sane A (2024) A self-regulated learning framework using generative AI and its application in CS educational intervention design. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education (SIGCSE 2024), Vol. 1* (pp. 1070-1076). Association for Computing Machinery. <https://doi.org/10.1145/3626252.3630828>
39. Puntambekar S, Hübscher R (2005) Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychol* 40(1):1-12. https://doi.org/10.1207/s15326985ep4001_1
40. Radovic S, Seidel N (2024) Self-regulated learning support in technology improved learning environments: A reliability analysis of the SRL-S rubric. *Int J Assess Tools Educ* 11(4):675-698. <https://doi.org/10.21449/ijate.1502786>

41. *Rahimi S, Shute V, Kuba R, Dai C-P, Yang X, Smith G, Fernández CA (2021) The use and effects of incentive systems on learning and performance in educational games. *Comput Educ* 165:104135. <https://doi.org/10.1016/j.compedu.2021.104135>
42. Ramadhanti D, Yanda DP (2021) Students' Metacognitive Awareness and Its Impact on Writing Skill. *Int J Lang Educ* 5(3):193. <https://doi.org/10.26858/ijole.v5i3.18978>
43. Rathika P, Ponni P, Anju P, Parthipan V, Yamunadevi S (2024) Developing an AI-Powered Interactive Virtual Tutor for improved Learning Experiences. *Int J Comput Experimental Sci Eng* 10(4). <https://doi.org/10.22399/ijcesen.782>
44. *Reinhold F, Hoch S, Werner B, Richter-Gebert J, Reiss K (2020) Learning fractions with and without educational technology: What matters for high-achieving and low-achieving students? *Learn Instruction* 65:101264. <https://doi.org/10.1016/j.learninstruc.2019.101264>
45. Reiser BJ (2004) Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *J Learn Sci* 13(3):273–304. https://doi.org/10.1207/s15327809jls1303_2
46. Ren Q, Li B, Chen Q, Zhang X (2023) Analysis of Academic Help-seeking Behavior in Online Learning. *2023 International Conference on Artificial Intelligence and Education (ICAIE)*, 15–19. <https://doi.org/10.1109/ICAIE56796.2023.00014>
47. Rigopouli K, Kotsifakos D, Psaromiligkos Y (2025) Vygotsky's Creativity Options and Ideas in 21st-Century Technology-improved Learning Design. *Educ Sci* 15(2):257. <https://doi.org/10.3390/educsci15020257>
48. Rizvi M (2023) Investigating AI-Powered Tutoring Systems that Adapt to Individual Student Needs, Providing Personalized Guidance and Assessments. *The Eurasia Proceedings of Educational and Social Sciences*, 31, 67–73. <https://doi.org/10.55549/epess.1381518>
49. Roll I, Winne PH (2015) Understanding, evaluating, and supporting self-regulated learning using learning analytics. *J Learn Analytics* 2(1):7–12. <https://doi.org/10.18608/jla.2015.21.2>
50. Roll I, Baker RSJ, d., Aleven V, Koedinger KR (2014) On the benefits of seeking (and avoiding) help in online problem-solving environments. *J Learn Sci* 23(4):537–560. <https://doi.org/10.1080/10508406.2014.883977>
51. Romi IM (2024) Digital skills impact on university students' academic performance: An empirical investigation. *Edelweiss Appl Sci Technol* 8(5):2126–2141. <https://doi.org/10.55214/25768484.v8i5.1965>
52. Ryan AM, Pintrich PR, Midgley C (2005) Avoiding seeking help in the classroom: Who and why? *Educational Psychol Rev* 13(2):93–114. <http://dx.doi.org/10.1023/A:1009013420053>

53. Samala AD, Zhai X, Bojic L, Zikic S, Aoki K (2024) An In-Depth Review of ChatGPT's Pros and Cons for Learning and Teaching in Education. *Int J Interact Mob Technol (iJIM)* 18(02):96–117. <https://doi.org/10.3991/ijim.v18i02.46509>
54. *Santhosh J, Dengel A, Ishimaru S (2024) Gaze-driven adaptive learning system with ChatGPT-generated summaries. *IEEE Access* 12:173714–173733. <https://doi.org/10.1109/ACCESS.2024.3503059>
55. Sardi J, Yuliana DF, Darmansyah D, Habibullah H, Yanto DTP, Eliza F, Candra O (2025) How Generative AI Influences Students' Self-Regulated Learning and Critical Thinking Skills? A Systematic Review. *Int J Eng Pedagogy (iJEP)* 15(1):94–108. <https://doi.org/10.3991/ijep.v15i1.53379>
56. Schunk DH, Zimmerman B (eds) (2011) *Handbook of Self-Regulation of Learning and Performance* (1st ed.). Routledge. <https://doi.org/10.4324/9780203839010>
57. *Serrano-Mendizábal M, Villalón R, Melero Á, Izquierdo-Magaldi B (2023) Effects of two computer-based interventions on reading comprehension: Does strategy instruction matter. *Comput Educ* 196:104727. <https://doi.org/10.1016/j.compedu.2023.104727>
58. Shabani K, Ebadi S, Khatib M (2010) Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development. *Engl Lang Teach* 3(4). <https://doi.org/10.5539/elt.v3n4p237>
59. Sharun K, Kumar R, Pawde AM, Dhama K, Pal A, Akash S, Banu SA (2023) ChatGPT and artificial hallucinations in stem cell research: assessing the accuracy of generated references – a preliminary study. *Annals Med Surg* 85(10):5275–5278. <https://doi.org/10.1097/ms9.0000000000001228>
60. Shen Y (2024) Examining the efficacies of instructor-designed instructional videos in flipped classrooms on student engagement and learning outcomes: An empirical study. *J Comput Assist Learn* 40(4):1791–1805. <https://doi.org/10.1111/jcal.12987>
61. Siddaway AP, Wood AM, Hedges L (2019) How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Ann Rev Psychol* 70:747–770. <https://doi.org/10.1146/annurev-psych-010418-102803>
62. Sidra S, Mason C (2025) Generative AI in Human-AI Collaboration: Validation of the Collaborative AI Literacy and Collaborative AI Metacognition Scales for Effective Use. *Int J Human-Computer Interact* 1–25. <https://doi.org/10.1080/10447318.2025.2543997>
63. Silva G, Godwin G, Jayanagara O (2024) The Impact of AI on Personalized Learning and Educational Analytics. *Int Trans Educ Technol (ITEE)* 3(1):36–46. <https://doi.org/10.33050/itee.v3i1.669>

64. Simms RC (2024) Work With ChatGPT, Not Against: 3 Teaching Strategies That Harness the Power of Artificial Intelligence. *Nurse Educ* 49(3):158-161. <https://doi.org/10.1097/nne.0000000000001634>
65. Singh B, Arora MK, Lal S, Kaunert C (2024) improving AI-Augmented Classrooms (pp. 99-130). *Igi Global*. <https://doi.org/10.4018/979-8-3693-7255-5.ch004>
66. Siu KWM, Yang Z, Zhao T, Zou J, Jiang Y, Zhang K (2025) Dynamic scaffolding: exploring the role of artificial intelligence in urban design education. *Front Urban Rural Plann* 3(1). <https://doi.org/10.1007/s44243-025-00060-7>
67. Stefaniak J, Moore S (2024) The Use of Generative AI to Support Inclusivity and Design Deliberation for Online Instruction. *Online Learn* 28(3). <https://doi.org/10.24059/olj.v28i3.4458>
68. Taub M, Azevedo R, Bradbury AE, Mudrick NV (2020) Self-regulation and reflection during game-based learning. In: Plass JL, Mayer RE, Homer BD (eds) *Handbook of game-based learning*. MIT Press, pp 239-262
69. *ter Beek M, Opdenakker M-C, Spijkerboer AW, Brummer L, Ozinga HW, Strijbos J-W (2019) Scaffolding expository history text reading: Effects on adolescents' comprehension, self-regulation, and motivation. *Learn Individual Differences* 74:101749. <https://doi.org/10.1016/j.lindif.2019.06.003>
70. Thomas DR, Aleven V, Koedinger KR, Fancsali SE, Lin J, Gupta S, Brunskill E, Norberg K, Gatz E, Branstetter L, Gurung A (2024) Improving Student Learning with Hybrid Human-AI Tutoring: A Three-Study Quasi-Experimental Investigation. 404-415. <https://doi.org/10.1145/3636555.3636896>
71. Valizadeh A, Moassefi M, Nakhostin-Ansari A, Hosseini Asl SH, Saghab Torbati M, Aghajani R, Ghorbani M, Z., Faghani S (2022) Abstract screening using the automated tool Rayyan: results of effectiveness in three diagnostic test accuracy systematic reviews. *BMC Med Res Methodol* 22(1). <https://doi.org/10.1186/s12874-022-01631-8>
72. Van De Pol J, Mercer N, Volman M (2018) Scaffolding Student Understanding in Small-Group Work: Students' Uptake of Teacher Support in Subsequent Small-Group Interaction. *J Learn Sci* 28(2):206-239. <https://doi.org/10.1080/10508406.2018.1522258>
73. Van de Pol J, Volman M, Beishuizen J (2010) Scaffolding in teacher-student interaction: A decade of research. *Educational Psychol Rev* 22:271-296. <https://doi.org/10.1007/s10648-010-9127-6>
74. Van Der Graaf J, Raković M, Fan Y, Lim L, Singh S, Bannert M, Gašević D, Molenaar I (2023) How to design and evaluate personalized scaffolds for self-regulated learning. *Metacognition Learn* 18(3):783-810. <https://doi.org/10.1007/s11409-023-09361-y>
75. Vanbelle S, Albert A (2009) Agreement between Two Independent Groups of Raters. *Psychometrika* 74(3):477-491. <https://doi.org/10.1007/s11336-009-9116-1>

76. VanLehn K (2006) The behavior of tutoring systems. *Int J Artif Intell Educ* 16(3):227-265
77. *Vicente-Yagüe-Jara MI, López-Martínez O, Navarro-Navarro V, Cuéllar-Santiago F (2023) Writing, creativity, and artificial intelligence. ChatGPT in the university context. *Comunicar* 31(77). <https://doi.org/10.3916/C77-2023-04>
78. *Vijayakumar S, Panwale SB (2025) Evaluating AI-Personalized Learning Interventions in Distance Education. *Int Rev Res Open Distrib Learn* 26(1):157-174. <https://doi.org/10.19173/irrodl.v26i1.7813>
79. Vilkova K, Shcheglova I (2020) Deconstructing self-regulated learning in MOOCs: In search of help-seeking mechanisms. *Educ Inform Technol* 26(1):17-33. <https://doi.org/10.1007/s10639-020-10244-x>
80. Villagrán I, Margozzini E, Miranda C, Fuentes J, Larrondo L, Iriarte Z, Hilliger I, Hernández R, Hurtado MT, Schuit G, Neyem A, Varas J (2024) improving Feedback Uptake and Self-Regulated Learning in Procedural Skills Training. *J Learn Analytics* 11(2):138-156. <https://doi.org/10.18608/jla.2024.8195>
81. Vogel F, Reiss K, Kollar I, Ufer S, Fischer F (2022) Adaptable scaffolding of mathematical argumentation skills: The role of self-regulation when scaffolded with CSCL scripts and heuristic worked examples. *Int J Computer-Supported Collaborative Learn* 17(1):39-64. <https://doi.org/10.1007/s11412-022-09363-z>
82. Vosniadou S (2020) Bridging Secondary and Higher Education. The Importance of Self-regulated Learning. *Eur Rev* 28(S1):S94-S103. <https://doi.org/10.1017/s1062798720000939>
83. Walker E, Koedinger KR, Rummel N (2011) Designing automated adaptive support to improve student helping behaviours in a peer tutoring activity. *Int J Computer-Supported Collaborative Learn* 6(2):279-306. <https://doi.org/10.1007/s11412-011-9111-2>
84. *Wang H, Huang T, Zhao Y, Hu S (2023) The impact of dashboard feedback type on learning effectiveness, focusing on learner differences. *Sustainability* 15(5):4474. <https://doi.org/10.3390/su15054474>
85. Wang Z, Nayfeh T, Murad MH, Tetzlaff J, O'Blenis P (2020) Error rates of human reviewers during abstract screening in systematic reviews. *PLoS ONE* 15(1):e0227742. <https://doi.org/10.1371/journal.pone.0227742>
86. Wen B, Liu J, Zulkernine F, Norel R, Chen H, Stappenbeck T (2024) Leveraging Large Language Models for Patient Engagement: The Power of Conversational AI in Digital Health. 104-113. <https://doi.org/10.1109/icdh62654.2024.00027>
87. Wilke J, Banzer W, Niederer D, Engeroff T, Vogt L, Nürnberger F, Krause F (2015) Appraising the methodological quality of cadaveric studies: validation of the QUACS scale. *J Anat* 226(5):440-446. <https://doi.org/10.1111/joa.12292>. Wiley

88. *Winkler R, Söllner M, Leimeister JM (2021) improving problem-solving skills with smart personal assistant technology. *Comput Educ* 165:104148. <https://doi.org/10.1016/j.compedu.2021.104148>
89. Wong J, Baars M, He M, De Koning BB, Paas F (2021) Facilitating goal setting and planning to improve online self-regulation of learning. *Comput Hum Behav* 124:106913. <https://doi.org/10.1016/j.chb.2021.106913>
90. Wood D, Bruner JS, Ross G (1976) The role of tutoring in problem solving. *J Child Psychol Psychiatry* 17(2):89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
91. *Wu C-H, Chen Y-S, Chen T-C (2018) An adaptive e-learning system for improving learning performance: Based on dynamic scaffolding theory. *EURASIA J Math Sci Technol Educ* 14(3):903–913. <https://doi.org/10.12973/ejmste/81061>
92. Wu G, Eastwood C, Sapiro N, Cheligeer C, Southern DA, Quan H, Xu Y (2024) Achieving high inter-rater reliability in establishing data labels: a retrospective chart review study. *BMJ Open Qual* 13(2):e002722. <https://doi.org/10.1136/bmj-oq-2023-002722>
93. Xi J, Lantolf JP (2020) Scaffolding and the zone of proximal development: A problematic relationship. *J Theory Social Behav* 51(1):25–48. <https://doi.org/10.1111/jtsb.12260>
94. Xiao N, Cai Z, Bu Y, Pei Y, Yuan C (2025) Transforming Education with Artificial Intelligence: A Comprehensive Review of Applications, Challenges, and Future Directions. *Int Theory Pract Humanit Social Sci* 2(1):337–356. <https://doi.org/10.70693/itphss.v2i1.211>
95. *Yan L, Martinez-Maldonado R, Jin Y, Echeverría V, Milesi M, Fan J, Zhao L, Alfredo R, Li X, Gašević D (2025) The effects of generative AI agents and scaffolding on improving students' comprehension of visual learning analytics. *Comput Educ* 234:105322. <https://doi.org/10.1016/j.compedu.2025.105322>
96. Yao Y, Shen J, Luo J, Li N, Liao X, Zhang Y (2024) Disagreements in risk of bias assessment for randomized controlled trials in hypertension-related Cochrane reviews. *Trials* 25(1). <https://doi.org/10.1186/s13063-024-08145-2>
97. Zaharuddin Z, Yu C, Yao G (2024) improving Student Engagement with AI-Driven Personalized Learning Systems. *Int Trans Educ Technol (ITEE)* 3(1):1–8. <https://doi.org/10.33050/itee.v3i1.662>
98. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F (2019) Systematic review of research on artificial intelligence applications in higher education – where are the educators? *Int J Educational Technol High Educ* 16:39. <https://doi.org/10.1186/s41239-019-0171-0>
99. Zimmerman BJ (2002) Becoming a self-regulated learner: An overview. *Theory Into Pract* 41(2):64–70. <https://doi.org/10.1207/s15430421tip41022>

Figures

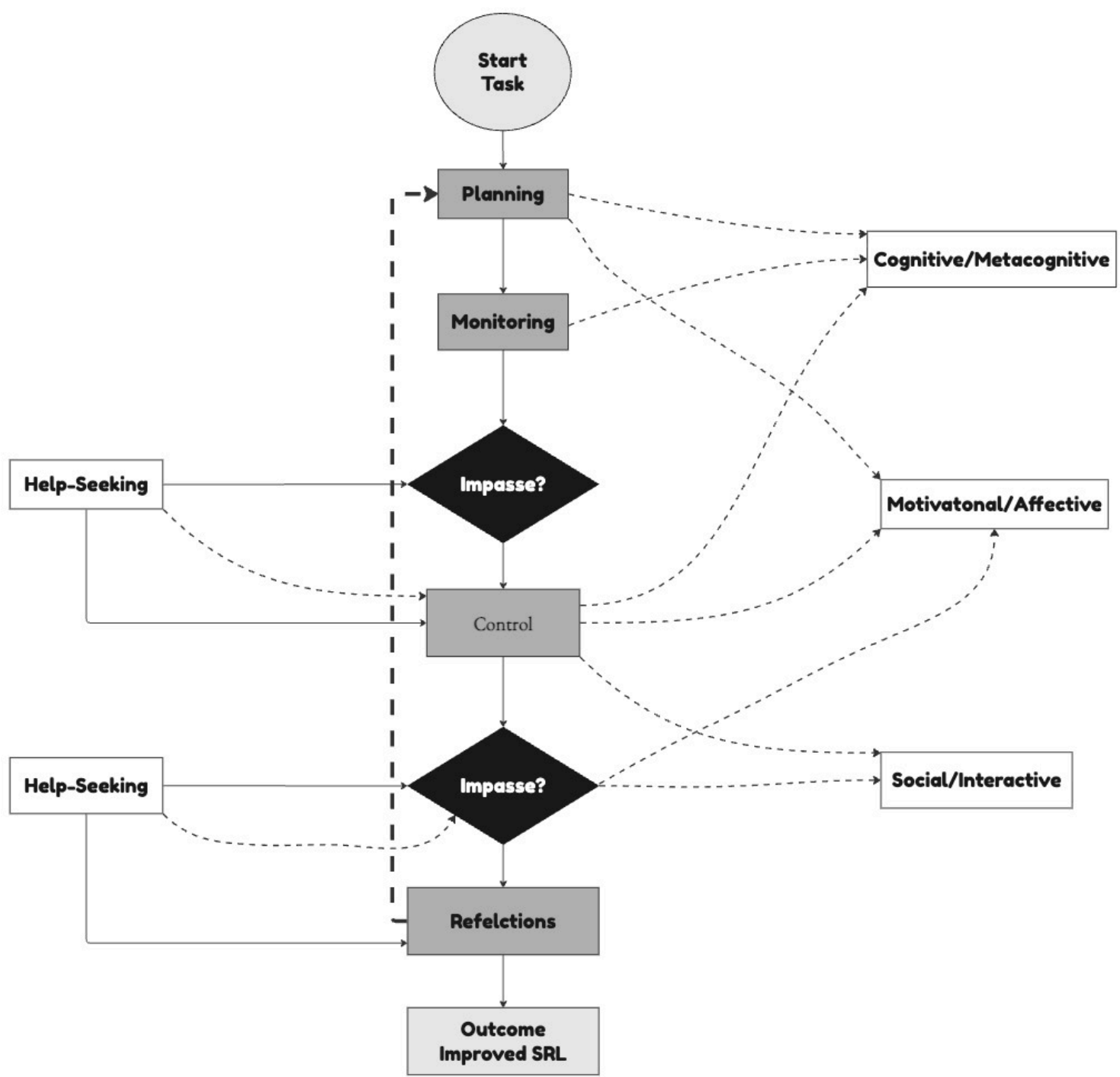


Figure 1

Integrating Help-Seeking into the Self-Regulated Learning Process: Theoretical Model

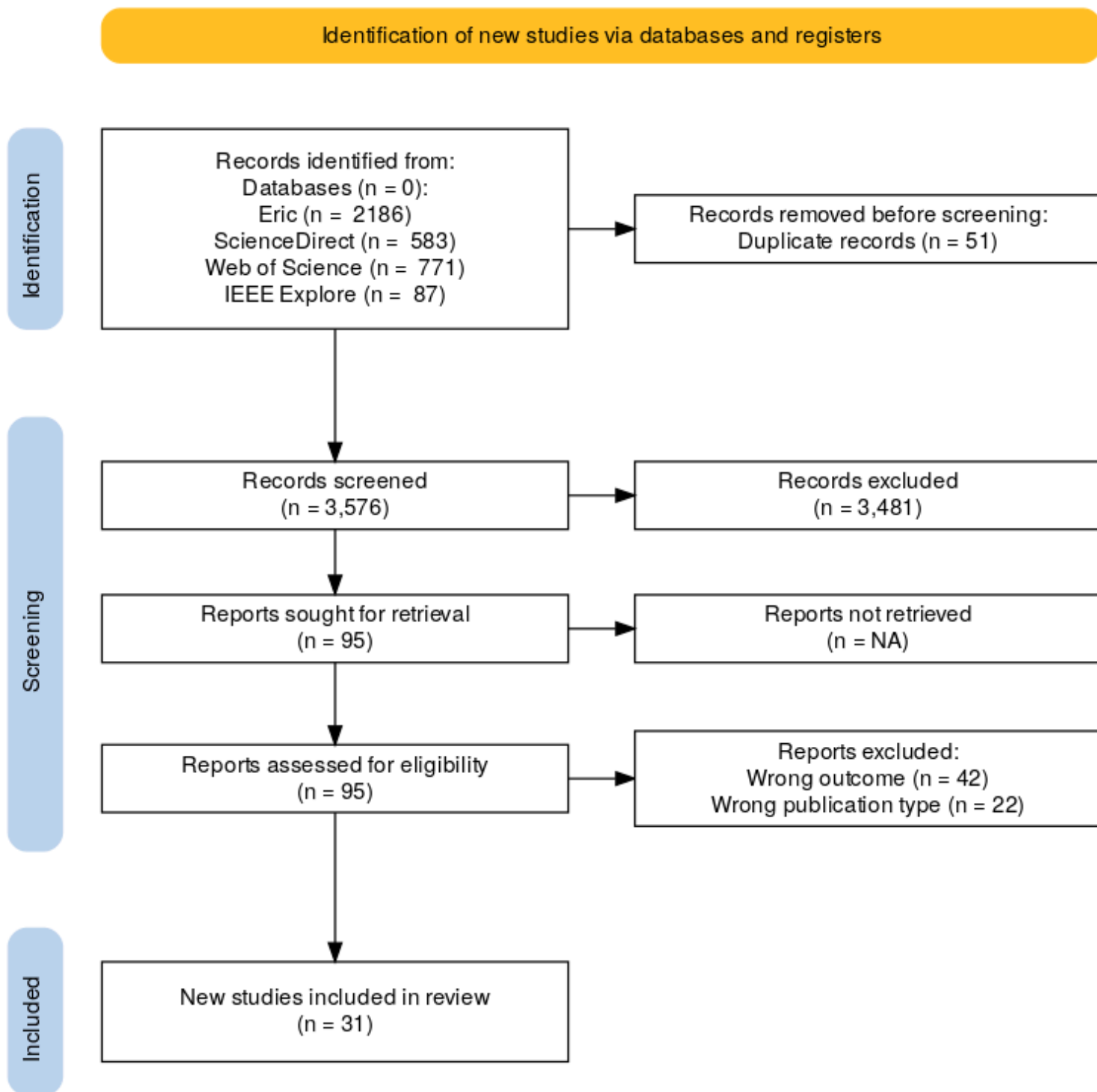


Figure 2

An overview of our search protocol based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) recommendation statement

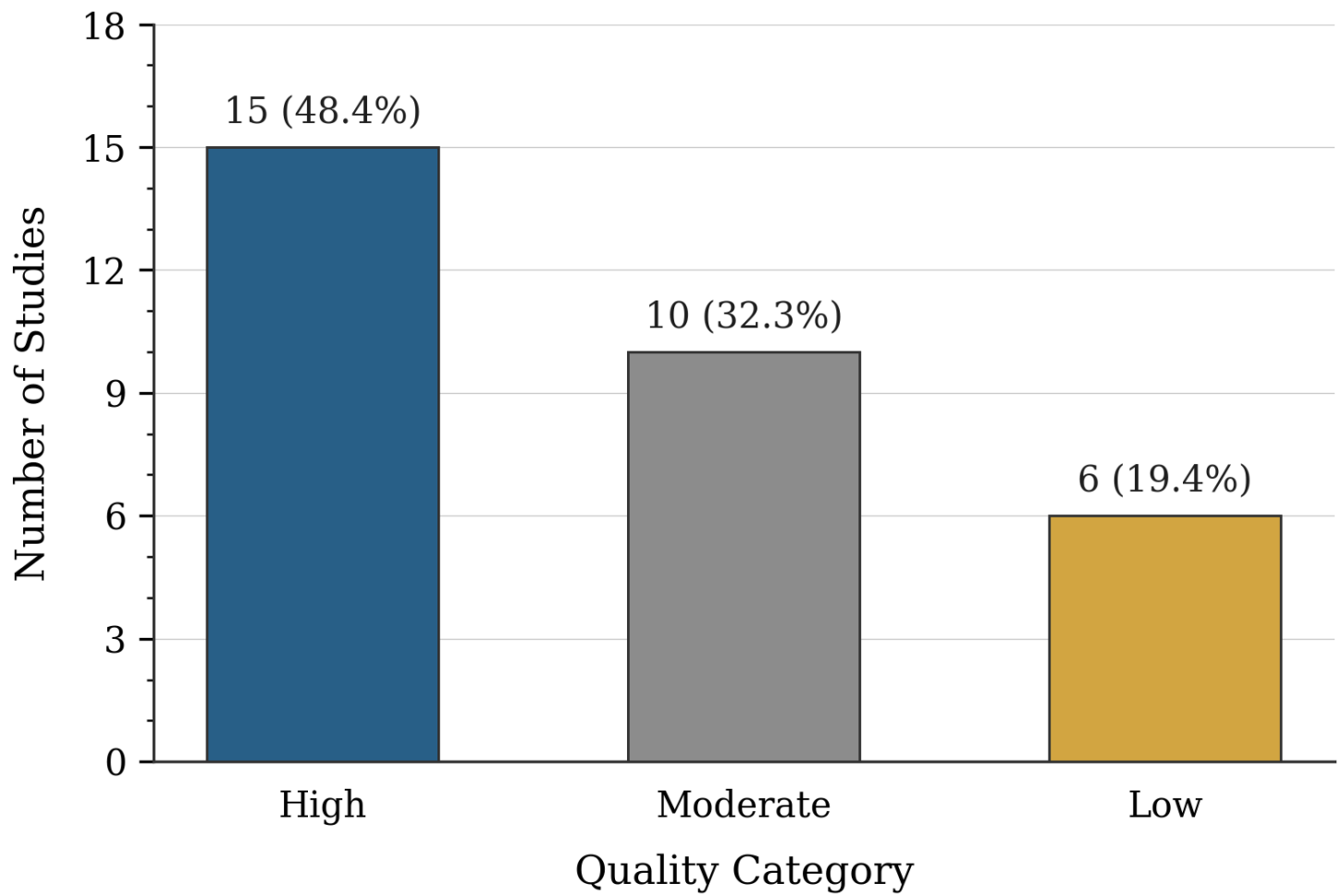


Figure 3

Quality Appraisal of Included Studies (n = 31)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AppendixA.docx](#)