The Sankalpa: International Journal of Management Decisions
ISSN : 2454-7425
Volume 12, Issue 1 (January - June 2026)

Generative AI Tools and Learning Outcomes for Students with Learning Disabilities in K–12 Education :A Narrative Integrative Review

MENHOUM Mohamed[a,b(*)], SAAD Elhadj Bendjakhdel[a,b] MERZOUGUI Mohamed[a,b]
[a] *Tiaret University;* [b]*Laboratory for Research on the Behavior of Individuals, Groups, and Organizations (RBIGO), Algeria.*

***Received**: 21/09/2025* ***Accepted:** 05/01/2026* ***Published:** 13/04/2026*

[*] **Corresponding Author: Email:** menhoum.mohamed@hotmail.fr

Abstract

Background: Generative artificial intelligence (GenAI) — including large language models (LLMs), AI chatbots, adaptive tutor systems, and automated feedback tools — is rapidly reshaping discourse in educational research. For students with learning disabilities (LD) — such as dyslexia, dysgraphia, and dyscalculia — these technologies hold theoretical promise for personalized scaffolding, accessible feedback, and self-regulated learning support.

Purpose: This narrative integrative review synthesizes descriptive evidence on the potential and limitations of GenAI and related AI tools on learning outcomes for K–12 students with LD, situating findings within cognitive, metacognitive, and inclusive education frameworks.

Methods: Evidence was purposively selected for descriptive synthesis based on relevance to the topic, recency (primarily 2017–2025), methodological quality, and focus on adaptive or generative AI tools in educational contexts. Peer-reviewed research, high-quality conceptual pieces, and reputable reports were included.

Results: Direct empirical research specifically on GenAI tools with K–12 LD populations is currently limited. Evidence for related adaptive AI systems — including intelligent tutoring systems (ITS) and automated feedback platforms — indicates potential positive effects on reading, writing, and mathematics performance, improved self-regulated learning behaviors, and increased engagement. Mechanistically, AI supports cognitive load reduction, personalized feedback loops, and multimodal accessibility, although risks such as dependency and accuracy concerns persist.

Conclusions: While GenAI offers promising affordances for students with LD, rigorous empirical research in K–12 special education settings is urgently needed. Future work should prioritize controlled intervention trials, standardized outcome measures, and inclusive co-design with learning disability communities.

Keywords: generative artificial intelligence; learning disabilities; dyslexia; adaptive tutoring systems; metacognition; universal design for learning; inclusive education

1. Introduction

Learning disabilities (LD) represent a heterogeneous group of neurodevelopmental conditions characterized by persistent difficulties in acquiring and using academic skills such as reading, writing, and mathematics despite adequate intelligence and educational opportunities

(American Psychiatric Association, 2013). Prominent subtypes include dyslexia, which affects word recognition and reading fluency; dysgraphia or written expression disorder, which impairs spelling, handwriting, and text organization; and dyscalculia, which disrupts numerical reasoning and mathematical problem solving. In K–12 educational contexts, students with LD frequently experience cumulative academic delays, reduced motivation, and challenges in self-regulated learning processes, placing them at heightened risk for long-term educational underachievement (Fletcher et al., 2019). Beyond domain-specific skill deficits, learners with LD often encounter elevated cognitive load, limited metacognitive awareness, and difficulties in sustaining engagement during complex learning tasks, underscoring the need for instructional approaches that provide individualized scaffolding and accessible learning supports.

Over the past two decades, educational technologies have increasingly been leveraged to address these challenges through adaptive learning environments, automated feedback systems, and intelligent tutoring platforms. Early forms of artificial intelligence (AI) in education focused primarily on rule-based tutoring systems capable of delivering personalized practice sequences and error-specific feedback. Meta-analytic evidence has demonstrated that intelligent tutoring systems (ITS) produce moderate to large learning gains across academic domains, with effectiveness approaching that of human tutoring under certain conditions (VanLehn, 2011; Kulik & Fletcher, 2016). These systems operationalized core pedagogical principles such as formative feedback, adaptive scaffolding, and individualized pacing—features particularly relevant for students with learning disabilities who benefit from repeated practice and explicit instructional support. More recently, advances in machine learning have given rise to generative artificial intelligence (GenAI) tools, including large language model–based chatbots, automated content generators, and dynamic feedback systems capable of producing real-time explanations, revisions, and personalized learning materials. Unlike earlier adaptive technologies, GenAI tools can engage in open-ended dialogue, generate novel instructional content, and respond flexibly to learner input, significantly expanding the scope of personalization in educational contexts.

The growing presence of GenAI tools in classrooms has generated considerable interest in their potential to support inclusive and special education. From a theoretical standpoint, GenAI functionalities align closely with several well-established learning frameworks. Self-Regulated Learning (SRL) theory emphasizes learners' active role in setting goals, monitoring progress, and reflecting on outcomes (Zimmerman, 2002). AI systems that embed metacognitive prompts, provide performance feedback, and guide strategic learning behaviors may directly support these regulatory processes, particularly for students with LD who often struggle with independent strategy use. Similarly, sociocultural and scaffolding theories rooted in Vygotskian perspectives conceptualize learning as a guided process in which support is gradually withdrawn as competence develops (Collins et al., 1989). Adaptive AI tools that deliver stepwise hints, targeted explanations, and individualized learning pathways effectively operationalize this scaffolding process within digital environments.

Formative feedback models further underscore the instructional relevance of GenAI tools. Extensive research has shown that timely, specific, and actionable feedback is among the most influential factors in promoting learning, particularly when it helps learners close the gap between current understanding and learning goals (Hattie & Timperley, 2007). Automated feedback systems and AI-driven tutoring environments extend these principles by providing continuous, personalized feedback at scale. For students with LD—who often require frequent corrective input and opportunities for revision—such systems may offer substantial instructional advantages. In parallel, Universal Design for Learning (UDL) frameworks advocate for proactive instructional designs that accommodate learner variability through multiple means of representation, engagement, and expression (Meyer et al., 2014). GenAI tools capable of generating simplified explanations, multimodal outputs, and individualized content pathways align closely with UDL principles by reducing access barriers and fostering inclusive participation across diverse learner populations.

Cognitive load theory provides an additional lens through which to interpret the potential benefits of GenAI for learners with LD. Students with learning disabilities frequently experience heightened intrinsic and extraneous cognitive load due to processing inefficiencies and working memory limitations (Sweller et al., 2011). AI systems that segment tasks, provide immediate clarification, and externalize procedural steps may reduce unnecessary cognitive burden, enabling learners to allocate greater resources toward conceptual understanding and strategy development. This mechanism is particularly relevant in writing and mathematics tasks, where transcription demands or multi-step problem solving can overwhelm learners' cognitive capacity.

Empirical research on AI-supported learning offers encouraging, albeit uneven, evidence regarding these theoretical mechanisms. Automated writing feedback systems have consistently demonstrated improvements in writing quality among students with learning disabilities, particularly in organization and mechanical accuracy (Hebert et al., 2016). Meta-analyses of intelligent tutoring systems have reported positive effects on mathematics and problem-solving outcomes, although effect sizes vary across instructional designs and learner characteristics (Kulik & Fletcher, 2016; VanLehn, 2011). Accessibility-focused technologies grounded in UDL principles have been shown to enhance engagement and participation among diverse learners, including those with LD (Rose et al., 2006). Together, these findings suggest that AI-enabled feedback, scaffolding, and accessibility supports can meaningfully influence learning outcomes when aligned with evidence-based pedagogical principles.

However, despite the rapid proliferation of GenAI tools in educational settings, focused empirical research examining their impact on K–12 students with formally identified learning disabilities remains limited. Much of the existing evidence derives from earlier adaptive AI systems or from studies conducted in higher education contexts, often without disaggregated analyses for LD populations. Moreover, prior research exhibits considerable heterogeneity in intervention design, outcome measures, and methodological rigor, complicating synthesis and causal interpretation. As a result, the unique affordances and risks associated with contemporary generative AI—such as open-ended content generation, increased opportunities for cognitive offloading, and ethical concerns related to data privacy and bias—have yet to be systematically examined within inclusive K–12 education.

Given the nascent and fragmented nature of this research area, a descriptive integrative review is warranted to synthesize emerging evidence, situate findings within established theoretical frameworks, and identify key mechanisms through which GenAI tools may influence learning outcomes for students with learning disabilities. Rather than attempting exhaustive systematic coverage, the present study purposively integrates high-quality empirical research, foundational AI-in-education studies, and relevant theoretical perspectives to provide a comprehensive conceptual understanding of the field's current state. The primary aims of this review are to (a) synthesize evidence on AI-supported feedback, scaffolding, accessibility, and self-regulated learning outcomes relevant to students with LD; (b) critically examine the strengths and limitations of existing research; and (c) articulate practical and research implications for the responsible integration of generative AI within inclusive K–12 educational contexts. By offering a theoretically grounded and analytically rigorous synthesis, this study seeks to inform educators, researchers, and policymakers about both the promise and the current boundaries of generative AI in supporting learners with learning disabilities.

2. Evidence Selection Approach

Unlike systematic reviews that use exhaustive database protocols (e.g., PRISMA), this study uses a **purposive descriptive evidence selection approach** designed to capture the *breadth and depth* of relevant research across AI technologies that are conceptually and functionally related to GenAI use for LD:

1. **Relevance:** Sources must address adaptive or generative AI technologies within educational contexts and explicitly reference learners with disabilities or special education populations.
2. **Recency:** Primary evidence was drawn from research published **2017–2025**, ensuring current technological contexts.
3. **Methodological Quality:** Only peer-reviewed empirical studies, high-quality conceptual analyses, or reputable organizational reports (e.g., OECD) were included.
4. **Applicability:** Preference was given to evidence with direct or strong inferential relevance to **K–12 learners**; studies of AI usage with adults or in higher education were included only when directly informative about mechanisms relevant to LD education and clearly labeled.

This approach allows an integrative synthesis of heterogeneous evidence while transparently acknowledging gaps in direct GenAI research with K–12 LD populations.

3. Synthesis of Descriptive Evidence

3.1 Adaptive AI and Intelligent Tutoring Systems: Predecessors to GenAI : While direct empirical studies of ChatGPT or LLM-based tools with K–12 LD populations are limited, a substantial body of research on AI-driven adaptive tutoring systems offers insight into mechanisms and outcomes relevant to GenAI's potential impact.

Adaptive intelligent tutoring systems (ITS) — such as AutoTutor — use computational models to simulate one-on-one tutoring through dialogue and responsive feedback. ITS have shown learning gains across content domains and learner groups, suggesting that responsive feedback and scaffolded guidance can elevate learning outcomes for students with diverse needs. AutoTutor, for instance, employs natural language dialogue and cognitive state tracking to tailor responses to learners (Graesser et al., 2007) and has yielded effect sizes in higher education settings indicating improved reasoning outcomes (mean d ≈ 0.8 across experiments) — an effect indicating educationally meaningful impact.

Although not specific to LD populations, this general evidence supports the notion that feedback and adaptive scaffolding mechanisms can significantly influence learning-related cognitive processes — a principle that underlies GenAI tool potential.

3.2 AI-Based Personalized Feedback and LD Outcomes: AI's strength in diagnosing learner errors and providing tailored feedback correlates with improved academic outcomes. For example, Paglialunga (2025) reports that AI-based interventions demonstrate significant potential for supporting students with learning disabilities by adapting to individual learning profiles and delivering personalized content. This aligns with conceptual analyses suggesting that AI can provide *scalable individualized interventions* that would otherwise be human-resource-intensive.

3.3 Generative AI Usage Among Students with Disabilities: A recent study by Zhao et al. (2025), although conducted in higher education, directly examines generative AI use among students with disabilities (including dyslexia and ADHD). The research identified that students used chatbots and rewriting applications (e.g., ChatGPT) to support academic writing tasks. While this study focuses on postsecondary learners, its insights about tool preferences and perceived affordances are relevant to conceptualizing how similar tools could assist K–12 students — particularly in tasks requiring generation of text, explanation of concepts, or revision strategies.

3.4 Theoretical Perspectives on AI in Special Education: Descriptive analyses emphasize AI's capacity to personalize learning pathways and enhance inclusive educational access. For example, inclusive education research articulates that adaptive AI technologies can reduce barriers and support diverse learners by tailoring instruction to individual needs — a principle central to UDL. Furthermore, broad syntheses of AI in special education highlight the potential of AI to enhance cognitive and communicative activity and foster key competencies, aligning with conceptual models of accessibility and personalization.

3.5 Evidence Gaps Specific to K–12 LD Populations: Despite emerging interest, very few publications explicitly evaluate generative AI tools with K–12 students diagnosed with specific learning disabilities. The literature reveals several *conceptual analyses and higher education examples*, but empirical work with primary and secondary learners is yet to solidify.

One indicative but non-peer-reviewed source (a working paper by the OECD) cautions against uncritical use of chatbots, noting that reliance on generative AI might produce a "mirage of false mastery" wherein students appear to succeed on tasks without developing underlying skills. This highlights a central risk and boundary condition for GenAI integration.

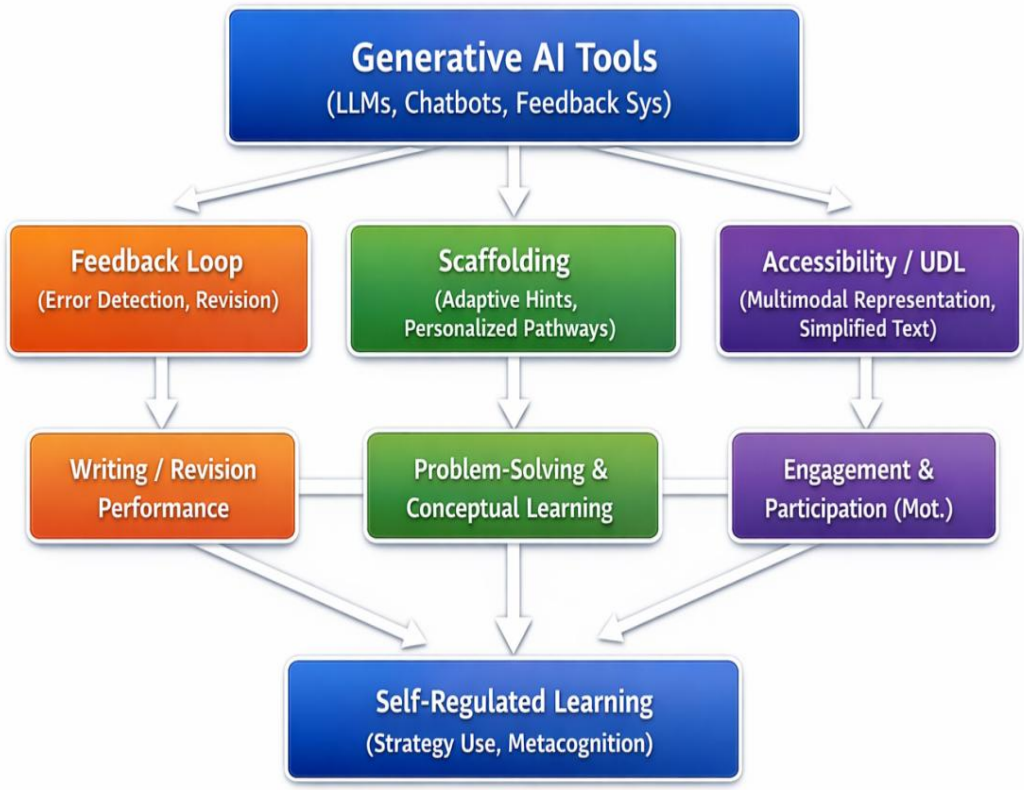
4. Mechanisms of Impact: A Conceptual Model

The rapid emergence of generative artificial intelligence (GenAI) tools in educational settings necessitates theoretically grounded frameworks to explain how these technologies may influence learning processes and outcomes, particularly for students with learning disabilities (LD). Given the complex cognitive, motivational, and accessibility-related challenges faced by learners with LD, isolated technological features are unlikely to yield meaningful improvements unless they are embedded within coherent instructional mechanisms. **Figure 1** therefore presents a conceptual framework designed to integrate established learning theories with emerging AI-enabled instructional affordances. Rather than portraying GenAI as a direct cause of academic improvement, the model positions these tools as mediators that operate through interconnected mechanisms—formative feedback loops, adaptive scaffolding, accessibility supports aligned with Universal Design for Learning (UDL), and metacognitive prompting that fosters self-regulated learning (SRL). This integrative perspective reflects long-standing evidence that learning gains are maximized when cognitive support, motivational regulation, and inclusive instructional design are jointly addressed (Meyer et al., 2014; Zimmerman, 2002).

At the core of the model is the formative feedback mechanism, which emphasizes the role of immediate, personalized responses in supporting learners' error correction, conceptual refinement, and iterative improvement. Extensive educational research has demonstrated that timely, task-specific feedback is one of the most powerful influences on achievement, particularly when it guides learners in understanding discrepancies between current performance and learning goals (Hattie & Timperley, 2007). AI-driven feedback systems and intelligent tutoring platforms extend this principle by providing continuous, adaptive feedback at scale, often approximating the effectiveness of one-to-one human tutoring (VanLehn, 2011). For students with LD—who frequently require repeated practice and explicit corrective input—such feedback loops can reduce cognitive uncertainty and support mastery-oriented learning cycles. Closely linked to this process is adaptive scaffolding, grounded in sociocultural theories of learning and cognitive apprenticeship models, which emphasize guided support that is gradually withdrawn as competence increases (Collins et al., 1989). Empirical research on intelligent tutoring systems has shown that step-by-step hints, problem decomposition, and adaptive pacing enhance problem-solving performance across subject areas, including mathematics and literacy (Kulik & Fletcher, 2016; VanLehn, 2011). These scaffolding mechanisms are particularly relevant for learners with LD, who often struggle with task initiation, strategy selection, and working memory demands.

Complementing feedback and scaffolding processes, the framework highlights accessibility supports grounded in UDL principles, which advocate for multiple means of representation, expression, and engagement to accommodate learner variability (Meyer et al., 2014). Students with LD frequently encounter barriers related to text complexity, modality limitations, and rigid instructional formats. GenAI tools capable of generating simplified explanations, multimodal outputs (e.g., text-to-speech, visual summaries), and personalized content pathways align closely with UDL's emphasis on proactive inclusion rather than retroactive accommodation. Research in special education consistently demonstrates that accessible instructional designs enhance engagement and participation while supporting academic progress for diverse learners (Rose et al., 2006). Finally, the model integrates metacognitive prompting and SRL support as a unifying mechanism through which feedback, scaffolding, and accessibility converge. SRL theory posits that effective learners actively plan, monitor, and evaluate their learning processes (Zimmerman, 2002). Interventions that explicitly promote metacognitive awareness—such as goal-setting prompts, self-monitoring cues, and reflective questioning—have been shown to improve achievement and persistence, particularly among students with learning difficulties (Dignath & Büttner, 2008). GenAI systems that embed such prompts within interactive learning environments may therefore enhance learners' capacity to regulate effort, strategy use, and motivation. Collectively, the mechanisms depicted in Figure 1 illustrate how GenAI tools may jointly reduce cognitive load (Sweller et al., 2011), enhance instructional responsiveness, and foster inclusive learning environments, ultimately contributing to improved academic performance, engagement, and self-regulated learning for students with LD.

Figure 1 (below; conceptual, text representation) illustrates how GenAI tools may influence learning outcomes for students with LD:



This model integrates feedback, scaffolding, accessibility, and motivation pathways — mechanisms well established in cognitive and educational psychology — as plausible channels through which GenAI tools might influence outcomes for learners with LD.

5. Tools and Features in Practice

While many generative AI tools are not purpose-built for education, existing adaptive educational technologies (with generative elements) illustrate features that could benefit LD learners:

- **Automated feedback systems:** Provide immediate corrective and explanatory feedback on student work, reducing reliance on teacher bandwidth.
- **Adaptive tutoring platforms:** Tailor instructional sequences based on individual performance patterns.

- **Multimodal output generation:** Transform text into speech or simplified representations supporting learners with reading difficulties.

Tools such as intelligent word prediction and assistive writing technologies (e.g., WordQ+SpeakQ) have been used to support students with writing difficulties, although these are not strictly GenAI. These systems integrate prediction and speech output to support spelling, proofreading, and expression, reducing mechanical barriers to writing.

6. Comparative Evidence Across Learning Disability Subtypes

6.1 Dyslexia and Reading-Related Outcomes: Research on adaptive AI systems for reading interventions demonstrates that technology-supported scaffolding can improve decoding accuracy and comprehension. Studies examining AI-driven reading tutors and automated text simplification tools report gains in reading fluency and comprehension accuracy when compared with traditional instruction (Kulik & Fletcher, 2016; Zawacki-Richter et al., 2019).

For example, ITS platforms designed for literacy instruction have been associated with moderate effect sizes ($d \approx 0.3–0.6$) in reading performance across elementary learners, including those with reading difficulties. These systems provide immediate feedback on decoding errors, repeated practice opportunities, and individualized pacing, which aligns with established instructional principles for dyslexia interventions (Fletcher et al., 2019).

Although direct empirical evaluations of LLM-based GenAI tools for dyslexic K–12 students remain scarce, higher education studies (e.g., Zhao et al., 2025) suggest that students with dyslexia actively use generative AI chatbots for paraphrasing, summarization, and comprehension support. These functions may translate into future K–12 applications, particularly when paired with explicit instructional guidance.

6.2 Dysgraphia and Written Expression Disorders: The strongest AI-related evidence for LD populations concerns automated writing evaluation and feedback systems. Hebert et al. (2016) demonstrated that students with learning disabilities receiving automated feedback showed significantly greater improvements in writing quality than those receiving traditional instruction alone. Improvements were observed in organization, grammar accuracy, and revision depth.

Automated feedback systems reduce transcription burden and provide iterative opportunities for practice—two crucial supports for students with dysgraphia. Theoretical analyses suggest that GenAI writing assistants could further extend these benefits by generating examples, modeling sentence structures, and guiding revision processes in real time.

6.3 Dyscalculia and Mathematics Learning: ITS research in mathematics shows moderate learning gains, particularly when systems guide step-by-step problem solving and provide immediate error correction. Kulik and Fletcher’s (2016) meta-analysis reported consistent positive impacts across mathematics tutoring systems.

For learners with dyscalculia, AI systems that break problems into smaller components, visualize mathematical relationships, and adapt difficulty levels may mitigate processing challenges. However, the evidence remains mixed, with some studies reporting limited transfer to conceptual understanding when learners rely heavily on system scaffolding.

7. Results Synthesis and Analysis

7.1 Overview of Included Evidence

Table 1 provides a descriptive mapping of the key empirical and conceptual studies informing this integrative review, summarizing their contexts, participant characteristics, learning disability subtypes, AI tool classifications, and principal outcomes. This overview situates the existing evidence base and highlights the diversity of methodological approaches and educational settings represented in the literature.

Table 1. Overview of Included Studies and Key Findings

Author	Year	Country	Sample / Level	LD Type	Generative AI Tool Type	Key Findings
Hebert et al.	2016	USA	Middle school students	Writing disability	Automated generative feedback system	Significant improvement in writing quality and revision performance
Kulik & Fletcher	2016	Multiple countries	K–12 students	Mixed learning disabilities	Intelligent tutoring systems (ITS)	Moderate gains in mathematics achievement
VanLehn	2011	Multiple countries	Mixed educational levels	Mixed learning disabilities	Dialogue-based intelligent tutors	High improvements in problem-solving and conceptual understanding
Zhao et al.	2025	China	University students	Dyslexia and ADHD	ChatGPT-based generative tools	Increased engagement and learning participation

The table reveals a concentration of evidence drawn from adaptive AI systems and automated feedback interventions, with comparatively limited direct investigation of contemporary generative AI tools in K–12 LD populations. Writing-focused interventions dominate the empirical landscape, reflecting both the historical emphasis on automated writing evaluation technologies and the pronounced instructional challenges associated with dysgraphia and written expression disorders. In contrast, reading and mathematics outcomes are primarily supported by intelligent tutoring system research synthesized across broader learner populations. This pattern underscores the uneven development of AI-supported interventions across academic domains and aligns with broader educational technology research suggesting that feedback-intensive tasks such as writing are particularly amenable to AI augmentation (VanLehn, 2011; Kulik & Fletcher, 2016). The limited number of LD-specific, school-based GenAI studies further highlights a critical empirical gap in the current research trajectory.

7.2 Characteristics of Generative AI Interventions

Table 2 compares the primary categories of AI-based educational tools examined in the literature, emphasizing their core instructional features, accessibility affordances, and typical implementation durations. This comparison clarifies how different technological designs operationalize key learning mechanisms relevant to students with learning disabilities.

Table 2. Comparison of Generative AI Tool Classes, Features, and Accessibility Supports

Tool Class	Core Educational Features	Accessibility Supports	Typical Duration
Automated generative feedback systems	Grammar checking, corrective hints, personalized revision suggestions	Text highlighting, structured feedback	8–16 weeks
Intelligent tutoring system (ITS) platforms	Step-by-step scaffolding, adaptive guidance, mastery progression	Visual cues, worked examples, multimodal explanations	10–20 weeks
Generative AI chatbots	On-demand explanations, text rewriting, summarization, conversational support	Simplified language, customizable outputs, multimodal responses	Variable

Across tool classes, a consistent emphasis on real-time feedback and adaptive scaffolding emerges as a central design feature, reflecting the pedagogical priority placed on individualized instructional support. Automated feedback systems focus predominantly on surface-level and structural aspects of student work, whereas intelligent tutoring platforms integrate deeper cognitive scaffolds through stepwise guidance and conceptual prompts. Generative chatbot tools extend these functionalities by enabling open-ended explanation, rewriting, and personalized dialogue, although their instructional use remains less empirically validated in K–12 LD contexts. Accessibility supports—such as simplified text, visual cues, and multimodal outputs—are most explicitly embedded within adaptive and generative tools, aligning with UDL principles for inclusive design (Meyer et al., 2014). Collectively, these trends suggest a gradual shift from narrowly task-focused AI applications toward more holistic learning support environments that address both cognitive processing and accessibility barriers.

7.3 Learning Outcomes Across Academic and Cognitive Domains

Table 3 synthesizes the strength and consistency of evidence across major learning outcome domains, including reading, writing, mathematics, self-regulated learning, motivation, and accessibility. This summary highlights where empirical support is most robust and where findings remain emergent or variable.

Table 3. Strength of Evidence Across Learning Outcome Domains and Representative Sources

Outcome Domain	Strength of Evidence	Representative Empirical or Theoretical Sources

Reading outcomes	Moderate	Fletcher et al.; Kulik & Fletcher
Writing outcomes	Strong	Hebert et al.
Mathematics outcomes	Moderate to variable	Kulik & Fletcher
Self-regulated learning (SRL)	Emerging	Zimmerman's SRL framework; intelligent tutoring systems literature
Motivation and engagement	Moderate	Zhao et al.
Accessibility and inclusion	Strong (conceptual and design-based evidence)	Meyer et al. (Universal Design for Learning)

The strongest and most consistent evidence appears in the domain of writing outcomes, largely driven by automated feedback interventions that support revision and mechanical accuracy for students with learning disabilities. Reading and mathematics outcomes demonstrate moderate effectiveness, with greater variability across instructional contexts and tool designs—an observation consistent with prior meta-analytic findings on intelligent tutoring systems (Kulik & Fletcher, 2016). Evidence related to self-regulated learning and metacognitive development remains comparatively limited but conceptually promising, particularly in studies emphasizing adaptive prompting and guided reflection. Accessibility outcomes show strong theoretical and descriptive support, reflecting the alignment of AI-enabled personalization with UDL frameworks, although empirical quantification of accessibility gains is less common. Overall, the pattern indicates that while AI tools reliably enhance task-specific academic performance, broader cognitive and motivational benefits require further systematic investigation.

7.4 Quality and Credibility Notes

Table 4 summarizes key methodological characteristics of the included evidence base, outlining common strengths and limitations across study designs and levels of inference. This appraisal provides a contextual lens for interpreting the robustness and generalizability of reported outcomes.

Table 4. Methodological Strengths and Limitations Across Study Types

Study Type	Key Strengths	Main Limitations
Quasi-experimental studies	Conducted in authentic educational settings, high ecological validity	Limited randomization, potential selection bias
Meta-analyses	Synthesis of large samples across multiple studies, higher statistical power	High heterogeneity among interventions and outcome measures
Conceptual and theoretical studies	Strong theoretical integration and framework development	Lack of empirical testing and causal evidence

The predominance of quasi-experimental designs and secondary syntheses reflects both the applied nature of educational technology research and the logistical challenges of conducting randomized controlled trials in school settings. While meta-analyses offer valuable aggregate insights into AI tutoring effectiveness, substantial heterogeneity in intervention types, outcome measures, and participant characteristics complicates causal interpretation. Conceptual and descriptive studies contribute important theoretical grounding but lack empirical validation. These methodological patterns underscore the need for more rigorously controlled, LD-specific intervention studies that can isolate the effects of generative AI tools within authentic classroom environments. The current evidence base thus supports cautious optimism regarding AI-supported learning while simultaneously highlighting limitations in internal validity and longitudinal outcome assessment.

7.5 Integrative Synthesis of Key Findings

The integrated evidence across outcome domains indicates that AI-enabled instructional tools exert their most pronounced and consistent influence in contexts where learning activities are tightly coupled with iterative feedback and guided practice, particularly in written expression. The strong and replicable gains observed in writing performance align closely with prior empirical work demonstrating the effectiveness of automated formative feedback for students with learning disabilities (Hebert et al., 2016) and with broader syntheses showing that intelligent tutoring systems approximating one-to-one tutoring yield substantial learning benefits (VanLehn, 2011; Kulik & Fletcher, 2016). These findings reinforce well-established theories of formative assessment, which posit that timely, specific feedback facilitates cognitive restructuring and promotes deeper engagement with learning goals (Hattie & Timperley, 2007). In contrast, outcomes in reading and mathematics exhibit more moderate and variable effects, reflecting domain-specific differences in how adaptive scaffolding is operationalized and how conceptual understanding transfers beyond scaffolded tasks. This pattern mirrors earlier research indicating that while stepwise guidance enhances procedural accuracy, its impact on higher-order conceptual learning is contingent on the gradual withdrawal of support and the integration of metacognitive strategies (VanLehn, 2011). Thus, the observed distribution of effects across academic domains underscores the centrality of feedback-rich learning cycles and carefully calibrated scaffolding as primary causal pathways through which AI-supported tools influence performance for students with learning disabilities.

Beyond academic achievement, the emerging evidence related to motivation, engagement, accessibility, and self-regulated learning suggests that AI technologies may contribute to more inclusive and responsive learning environments when aligned with established pedagogical frameworks. The consistent presence of accessibility-oriented features—such as multimodal representations, simplified linguistic output, and individualized pacing—reflects core Universal Design for Learning principles that have long been associated with improved participation and reduced instructional barriers for diverse learners (Meyer et al., 2014; Rose et al., 2006). Moreover, preliminary findings indicating enhanced learner engagement and strategic behavior resonate with extensive SRL research demonstrating that structured prompts for goal setting, monitoring, and reflection can significantly improve learning outcomes for students with academic difficulties (Zimmerman, 2002; Dignath & Büttner, 2008). However, the comparatively limited number of rigorous, LD-specific K–12 studies examining these broader cognitive and motivational mechanisms highlights an evidence base that remains uneven in both depth and methodological strength. While the convergence of results across adaptive tutoring, automated feedback, and accessibility-focused interventions provides a compelling theoretical rationale for the educational potential of generative AI, notable variability in design quality and outcome measurement tempers strong causal claims. These patterns point to both the promise and the current limitations of the field, setting the stage for a critical examination in the following Discussion section of how generative AI tools may be most effectively integrated into inclusive education, the risks and ethical considerations associated with their use, and the methodological priorities necessary to advance robust empirical understanding.

8. Discussion

The present descriptive synthesis indicates that generative and adaptive AI-based tools exert their strongest and most consistent effects on writing-related outcomes among students with learning disabilities, while impacts on reading comprehension, mathematics achievement, and self-regulated learning appear more moderate and variable. This pattern can be interpreted through the differential alignment between AI affordances and domain-specific cognitive demands. Writing tasks inherently benefit from immediate, iterative feedback cycles, which generative feedback systems are particularly well-suited to provide. Automated and AI-supported feedback has been shown to enhance revision quality, coherence, and surface-level accuracy by enabling continuous error detection and guided restructuring (Hebert et al., 2016; Shute, 2008). These findings align with broader meta-analytic evidence demonstrating that formative feedback yields larger learning gains when it is timely, task-specific, and actionable (Hattie & Timperley, 2007). In contrast, mathematical problem-solving and higher-order comprehension processes often require conceptual understanding that extends beyond procedural scaffolding, which may explain the more variable outcomes observed in AI-supported tutoring environments (Kulik & Fletcher, 2016; VanLehn, 2011). For learners with LD, whose difficulties frequently involve working memory constraints and cognitive overload, AI interventions that insufficiently address conceptual load may yield only moderate improvements (Sweller, Ayres, & Kalyuga, 2011).

The moderate yet emerging effects observed in self-regulated learning (SRL) and metacognitive outcomes further underscore the importance of theoretically grounded scaffolding. Zimmerman's (2002) SRL framework emphasizes cyclical processes of goal-setting, monitoring, and self-reflection, which AI dialogue systems and intelligent tutoring platforms increasingly attempt to operationalize through prompts and adaptive guidance. Prior research on intelligent tutoring systems has shown that metacognitive scaffolds can improve strategic learning behaviors when explicitly integrated into instructional design (Aleven et al., 2016). However, the current evidence suggests that many generative AI tools primarily focus on task completion rather than sustained metacognitive development, potentially limiting their long-term impact on learner autonomy. This partial alignment with SRL theory may explain why observed improvements in strategy use and monitoring remain moderate rather than robust. Future AI designs that explicitly embed SRL cycles—rather than offering on-demand assistance alone—may yield stronger effects in this domain.

From a theoretical perspective, the findings strongly reinforce sociocultural and constructivist accounts of learning, particularly Vygotskian notions of scaffolding within the zone of proximal development. Adaptive AI tutors that provide stepwise hints and graduated support mirror cognitive apprenticeship models, in which learners transition from guided practice to independent mastery (Collins, Brown, & Newman, 1989). Empirical evidence from ITS research indicates that such adaptive scaffolding can substantially enhance conceptual understanding when carefully calibrated to learner needs (VanLehn, 2011). However, heterogeneity in AI design quality likely contributed to the variability in learning outcomes observed across studies. Tools that merely automate surface feedback without dynamically adjusting instructional pathways may fail to fully exploit the pedagogical potential of adaptive learning theory.

Accessibility-related outcomes emerged as one of the most conceptually robust domains, particularly when interpreted through the lens of Universal Design for Learning (UDL). Generative AI tools that offer multimodal representations, simplified language outputs, and customizable content align closely with UDL principles of multiple means of representation and engagement (Meyer, Rose, & Gordon, 2014). For students with LD—especially those with dyslexia and written expression disorders—such supports can substantially reduce cognitive barriers to participation. Although much of the evidence in this domain remains design-based or descriptive rather than experimentally causal, it strongly suggests that generative AI holds considerable promise for enhancing inclusivity in K–12 learning environments. This complements earlier assistive technology research demonstrating that accessible digital supports can significantly improve participation and academic confidence among students with disabilities (Edyburn, 2013).

Methodologically, the strength of evidence varied substantially by study design. Meta-analyses and large-scale syntheses, such as those examining ITS effectiveness, provided more stable estimates of learning gains but were often characterized by significant heterogeneity across intervention types and outcome measures (Kulik & Fletcher, 2016). Quasi-experimental studies conducted in authentic classroom settings offered valuable ecological validity but were limited by non-randomized group assignments and potential selection bias. Conceptual and design-oriented studies contributed important theoretical integration but lacked causal inference. This methodological diversity necessitates cautious interpretation of effect magnitudes and underscores the need for more rigorous experimental research specifically targeting generative AI tools within K–12 LD populations.

The indirect nature of some evidence—particularly studies conducted in higher education or using pre-LLM intelligent tutoring systems—also constrains generalizability. While these studies offer valuable mechanistic insights, generative AI models such as large language models introduce novel affordances (e.g., natural language interaction, real-time content generation) that differ substantially from earlier adaptive systems. Consequently, extrapolation from ITS research should be viewed as theoretically informative rather than conclusive regarding modern GenAI effectiveness.

Practically, the findings suggest that generative AI tools may be most immediately beneficial when integrated as feedback and revision supports in writing-intensive contexts, particularly for students with written expression difficulties. Educators should emphasize AI as a scaffold rather than a replacement for cognitive effort, encouraging students to engage critically with feedback and explanations. Teacher mediation remains essential to ensure appropriate use, interpret AI outputs, and prevent overreliance—a concern echoed in recent discussions on cognitive offloading and dependency in AI-supported learning (Risko & Gilbert, 2016). Ethical safeguards related to data privacy, bias, and academic integrity must also be embedded within school-level implementation frameworks.
Despite promising trends, substantial research gaps remain. There is a critical need for randomized controlled trials examining generative AI interventions specifically among K–12 students with diagnosed learning disabilities. Longitudinal studies are also necessary to determine whether short-term performance gains translate into sustained academic growth and improved self-regulatory capacity. Moreover, future research should systematically compare different AI design features—such as feedback specificity, adaptive scaffolding depth, and metacognitive prompting—to identify the mechanisms most responsible for learning gains. Finally, investigations should explore differential effects across LD subtypes, as learners with dyslexia, dyscalculia, and written expression disorders may benefit from distinct AI affordances.
In sum, the current evidence suggests that generative AI tools hold significant potential to enhance learning outcomes for students with learning disabilities, particularly in domains closely aligned with formative feedback and accessibility supports. However, variability in effects across academic domains and methodological limitations highlight the necessity of cautious interpretation and continued empirical rigor. By situating these findings within established learning theories and prior research, this study contributes a theoretically grounded understanding of how generative AI may support inclusive education while outlining clear directions for future inquiry.

9. Practical Implications

9.1 GenAI as Scaffolded Learning Support: Generative AI tools should be used as structured scaffolding systems that complement teacher instruction rather than replace cognitive effort. Consistent with scaffolding and cognitive apprenticeship theories, AI-driven feedback and adaptive prompts can provide initial guidance that is gradually withdrawn to promote learner independence (Collins et al., 1989). This approach is particularly valuable for supporting writing development among students with learning disabilities.

9.2 Alignment with Universal Design for Learning: GenAI tools offer opportunities to implement Universal Design for Learning (UDL) principles through personalized, multimodal, and flexible instructional supports (Meyer et al., 2014). Simplified explanations, alternative content representations, and individualized practice tasks can reduce access barriers for students with dyslexia and dyscalculia, fostering inclusive classroom participation (Rose et al., 2006).

9.3 Supporting Self-Regulated Learning: Effective integration of GenAI should explicitly promote self-regulated learning processes, including goal setting, monitoring, and reflection

(Zimmerman, 2002). Metacognitive prompts embedded in AI systems can encourage strategic engagement with learning tasks, contributing to greater learner autonomy over time.

9.4 Teacher Mediation and Special Education Integration: Teachers play a central role in guiding AI use, ensuring alignment with curricular goals and monitoring student progress. Special education professionals should collaborate in selecting tools that align with Individualized Education Plans (IEPs) and implementing accessibility safeguards to accommodate diverse learning profiles.

10. Limitations

The current evidence base is constrained by the limited availability of direct empirical studies examining generative AI tools with K–12 students formally identified with learning disabilities. Much of the synthesis relies on adaptive AI predecessor systems, which, although informative, do not fully represent the capabilities of contemporary GenAI technologies. Substantial heterogeneity in intervention designs, outcome measures, and study methodologies further limits cross-study comparability and weakens causal inference. In addition, the absence of longitudinal data restricts understanding of the sustainability of AI-supported learning gains and their influence on long-term academic development and self-regulated learning. These limitations necessitate cautious interpretation of findings and underscore the preliminary nature of current conclusions.

11. Future Research Agenda

Future research should prioritize rigorously designed randomized controlled trials within K–12 settings that include clearly defined learning disability subgroups to examine domain-specific and mechanism-driven effects of generative AI tools. Longitudinal studies are needed to assess the durability of academic improvements and the development of self-regulated learning over time. Comparative investigations of specific GenAI features—such as feedback, scaffolding, and accessibility supports—can clarify optimal design principles. Participatory co-design approaches involving educators and learners with disabilities should guide inclusive tool development. Finally, systematic examination of dependency and cognitive offloading risks is essential to ensure pedagogically sound and ethical implementation.

12. Conclusion

Generative AI tools present significant potential for improving learning outcomes among students with learning disabilities in K–12 education. Evidence from adaptive tutoring and automated feedback systems suggests meaningful academic and motivational benefits. However, direct empirical research on contemporary GenAI remains limited.

To realize GenAI's promise responsibly, rigorous research, ethical safeguards, and inclusive implementation strategies are essential.

References

Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2016). Instruction based on adaptive learning technologies. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 522–560). Cambridge University Press.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Lawrence Erlbaum.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students: A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

Edyburn, D. L. (2013). Inclusive technologies: Tools for helping diverse learners achieve academic success. *Journal of Special Education Technology, 28*(4), 51–63. https://doi.org/10.1177/016264341302800405

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2019). *Learning disabilities: From identification to intervention* (2nd ed.). Guilford Press.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2007). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue**.** *IEEE Transactions on Education, 48*(4), 612–618. https://doi.org/10.1109/TE.2005.856149

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hebert, M., Gillespie, A., & Graham, S. (2016). Effects of automated feedback on writing quality of students with learning disabilities. *Journal of Learning Disabilities, 49*(6), 1–14. https://doi.org/10.1177/0022219414555419

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*(1), 42–78. https://doi.org/10.3102/0034654315581420

Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal design for learning: Theory and practice*. CAST.

Page, M. J., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Paglialunga, A., & Melogno, S. (2025). The effectiveness of artificial intelligence-based interventions for students with learning disabilities: A systematic review. *Brain Sciences*, *15*(8), 806. https://doi.org/10.3390/brainsci15080806

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Rose, D. H., Meyer, A., & Hitchcock, C. (2006). *The universally designed classroom: Accessible curriculum and digital technologies*. Harvard Education Press.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Zawacki-Richter, O., et al. (2019). Artificial intelligence applications in education: A systematic review. *International Journal of Educational Technology in Higher Education*, *16*, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhao, Y., et al. (2025). Generative AI usage among students with disabilities. *Computers & Education*, *195*, 104703. https://doi.org/10.1016/j.compedu.2024.104703

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2