Artificial Intelligence in Enhancing Teaching and Learning: A Systematic Review

Rajneesh Kumar Srivastava
Bhavdiya Educational Institute
Ayodhya, India

Abstract- **Artificial intelligence (AI) is rapidly reshaping the educational landscape, offering unprecedented opportunities to individualize instruction, automate administrative burdens, and generate data-driven insights into the learning process. This review paper systematically examines the theoretical foundations and empirical evidence for AI applications across six domains of educational practice: intelligent tutoring systems (ITS), adaptive learning platforms, automated assessment and feedback, natural language processing for literacy development, AI-driven learning analytics, and large language model (LLM) integration in classrooms. Drawing on peer-reviewed literature published between 2015 and 2025, we critically appraise evidence for learning outcome improvements, analyse the pedagogical frameworks underpinning AI-enhanced instruction, and examine equity, privacy, and ethical dimensions of AI deployment in educational settings. Our synthesis reveals substantial empirical support for AI-driven personalization improving student achievement by 0.3–0.8 standard deviations compared to conventional instruction in controlled studies, while also identifying persistent challenges including data bias, digital access inequity, teacher readiness, and the risk of over-automation of formative human relationships central to effective pedagogy. We conclude with a research agenda and policy recommendations for evidence-based AI integration in K-12, higher education, and lifelong learning contexts.**

Keywords: Artificial intelligence, education, intelligent tutoring systems, adaptive learning, automated feedback, learning analytics, large language models, personalized learning, educational technology, formative assessment

1. INTRODUCTION

Education systems worldwide face a constellation of persistent challenges: growing classroom diversity that strains the capacity of a single teacher to differentiate instruction for every learner; an assessment infrastructure that produces infrequent, summative snapshots of learning rather than actionable, real-time diagnostic information; and an expanding body of pedagogical research whose insights struggle to penetrate routine classroom practice at scale. Artificial intelligence presents a technological framework that, if thoughtfully designed and deployed, can address each of these challenges by enabling instructional systems to observe learner behaviour at high temporal resolution, infer cognitive and affective states from that behaviour, and adapt instructional inputs dynamically in response [1].

The intersection of AI and education is not new. Intelligent tutoring systems have existed since the 1970s, with SCHOLAR—a geography tutoring system developed by Carbonell [2]—representing an early landmark. However, the confluence of three developments in the past decade has dramatically accelerated the pace and scope of AI integration in education: the availability of large-scale digital learning data from learning management systems, online course platforms, and ed-tech applications; advances in machine learning, natural language processing, and deep learning that enable more sophisticated modelling of learner behavior; and the emergence of powerful generative AI systems, particularly large language models, that can engage in open-ended instructional dialogue and generate personalized educational content [3].

The global AI in education market was valued at approximately USD 4 billion in 2022 and is projected to grow to USD 30 billion by 2032, reflecting substantial investment in AI-enabled learning platforms, assessment tools, and administrative automation systems [4]. This rapid growth makes critical, evidence-based appraisal of AI educational applications both timely and essential. This review is organized as follows. Section II examines intelligent tutoring systems. Section III reviews adaptive learning platforms. Section IV covers automated assessment and feedback. Section V examines NLP for literacy development. Section VI discusses learning analytics. Section VII addresses LLM integration. Section VIII covers challenges and ethical considerations. Section IX proposes future research directions, followed by conclusions in Section X.

2. INTELLIGENT TUTORING SYSTEMS

A. Architecture and Theoretical Foundations

Intelligent tutoring systems (ITS) are computer-based learning environments that provide individualized instruction and feedback without requiring real-time human intervention. The canonical ITS architecture, formalized by Woolf [5], comprises four components: a domain model encoding the knowledge structure of the subject matter; a learner model tracking the student's current knowledge state, misconceptions, and learning trajectory; a pedagogical model specifying instructional strategies; and an interface through which learner and system interact. Modern ITS implementations augment this architecture with machine learning components that enable learner models to update probabilistically as evidence of student knowledge accumulates, typically using Bayesian knowledge tracing (BKT) or deep knowledge tracing (DKT) frameworks [6].

Bayesian knowledge tracing, introduced by Corbett and Anderson [7], models student knowledge as a hidden Markov model in which the probability of a student knowing a skill is updated based on observed correct and incorrect responses. Deep knowledge tracing, proposed by Piech et al. [6], replaces the latent variable model with a recurrent neural network trained on large datasets of student interaction sequences, achieving significantly improved prediction accuracy of future student performance on held-out problems. These knowledge tracing models form the inferential backbone of adaptive ITS, enabling the system to select instructional content at appropriate difficulty levels and intervene with targeted remediation when knowledge gaps are detected.

B. Empirical Evidence for Learning Outcomes

Meta-analytic evidence for ITS effectiveness is among the strongest in educational technology research. VanLehn's [8] seminal meta-analysis of 26 ITS evaluations found a mean effect size of d = 0.76 compared to conventional classroom instruction—substantially larger than the typical effect size for educational interventions—and comparable to one-on-one human tutoring, consistent with Bloom's [9] "2 sigma" finding. Subsequent meta-analyses have refined this estimate: Kulik and Fletcher [10] reported a mean effect size of d = 0.66 across 50 ITS studies, with stronger effects for mathematics and science subjects and for students with lower prior knowledge. Carnegie Learning's MATHia platform, one of the most rigorously evaluated ITS implementations, demonstrated statistically significant improvements in standardized mathematics assessment scores in a randomized controlled trial spanning 147 schools [11].

3. ADAPTIVE LEARNING PLATFORMS

A. Personalized Learning Pathways

Adaptive learning platforms extend the ITS concept to broader curriculum delivery, dynamically sequencing and pacing instructional content based on continuously updated models of individual learner proficiency. Platforms such as Khan Academy, Knewton Alta, and Smart Sparrow employ item response theory (IRT) and machine learning models to estimate learner ability levels and select instructional items that maximize expected learning gain—a concept formalized as the "zone of proximal development" in Vygotskian learning theory [12]. The alignment between adaptive learning system design and constructivist and zone-of-proximal-development theories provides a theoretical rationale for why personalized sequencing should enhance learning that extends beyond empirical results.

Dziuban et al. [13] conducted a large-scale evaluation of adaptive learning in higher education mathematics courses, finding that students in adaptive courses achieved pass rates 9–18 percentage points higher than comparison groups using traditional instruction, with the largest gains concentrated among historically underserved student populations. A randomized controlled trial of Knewton's adaptive platform in introductory statistics found that students in the adaptive condition spent 16% less time to reach equivalent proficiency levels compared to linear instruction [14], suggesting efficiency gains alongside outcome improvements. The reduction in time-to-mastery has significant implications for educational equity, enabling students with caregiving or employment obligations to complete coursework within constrained time budgets.

B. Affective and Motivational Adaptation

An emerging dimension of adaptive learning research concerns affective computing—the detection of emotional states such as frustration, boredom, confusion, and engagement from behavioural signals including response latency, error patterns, help-seeking behaviour, and, where available, facial expression and physiological data. Baker et al. [15] demonstrated that boredom during ITS interaction is a significant predictor of long-term disengagement and drop-out, motivating the development of affect-aware tutoring systems capable of modifying motivational strategies—introducing challenge variation, offering encouragement, or altering task difficulty—in response to detected affective states. D'Mello and Graesser [16] developed AutoTutor with affective sensitivity, showing that an emotionally aware version of the tutoring system produced significantly larger learning gains than an affect-blind version, particularly for students prone to frustration.

4. AUTOMATED ASSESSMENT AND FEEDBACK

A. Automated Essay Scoring

Automated essay scoring (AES) systems apply natural language processing and machine learning to assign holistic or analytic scores to student-written texts, with the dual goals of reducing grading burden and providing more frequent, timely feedback than human grading workflows allow. The earliest AES systems, including Project Essay Grade (PEG) developed by Page [17] in 1966, relied on surface-level proxy features such as word count, sentence length, and lexical frequency. Contemporary AES systems leverage transformer-based language models, including fine-tuned versions of BERT and RoBERTa, to capture semantic coherence, argument structure, and discourse-level qualities [18]. Validus Studies meta-analyses report AES-human score agreement Pearson correlations of $r = 0.80–0.95$ across major commercial systems including e-rater (ETS), Turnitin's WritingMatch, and Pearson's Intelligent Essay Assessor [19].

A persistent concern with AES systems is their susceptibility to adversarial inputs—texts that achieve high automated scores through strategic manipulation of surface features without conveying genuine meaning or demonstrating substantive understanding. Perelman [20] demonstrated that a deliberately nonsensical essay composed of high-frequency academic vocabulary could receive above-average scores from commercial AES systems, highlighting the divergence between machine-scored proxies for quality and the holistic human judgments the systems are intended to emulate. Addressing this limitation requires AES systems that evaluate deeper semantic coherence and factual accuracy, areas where recent LLM-based approaches show promise but require further validation.

B. Formative Feedback and Automated Writing Assistance

Beyond summative scoring, AI is increasingly deployed to deliver formative, process-oriented feedback during the writing process rather than after submission. Systems such as Grammarly, Turnitin's Draft Coach, and Microsoft Editor provide real-time feedback on grammar, style, clarity, and citation practices, enabling students to iteratively revise their work with AI-mediated scaffolding. Rankin et al. [21] conducted a quasi-experimental study of an AI writing feedback system in secondary school English courses, finding that students who received automated process feedback during drafting produced significantly higher-quality final essays ($d = 0.52$) compared to students receiving only summative teacher feedback after submission, and developed more sophisticated self-revision strategies over the academic year.

5. NATURAL LANGUAGE PROCESSING FOR LITERACY DEVELOPMENT

A. Reading Comprehension and Vocabulary

Natural language processing technologies enable a range of literacy support applications that adapt to individual student reading levels, vocabulary knowledge, and comprehension profiles. Automated readability analysis systems can estimate the reading complexity of texts across multiple dimensions—syntactic complexity, lexical frequency, conceptual density—enabling platforms to dynamically adjust text difficulty or provide targeted vocabulary scaffolding for individual readers [22]. Reading comprehension question generation systems, powered by transformer models fine-tuned on educational datasets, can automatically generate inferential and evaluative questions from arbitrary texts, reducing the time required for teachers to design differentiated reading materials [23].

B. Intelligent Language Learning Systems

AI has transformed language learning applications, with platforms such as Duolingo deploying machine learning models to optimize lesson sequencing, predict the optimal timing of vocabulary review using spaced repetition algorithms, and provide adaptive pronunciation feedback using automatic speech recognition (ASR) [24]. Settles et al. [25] described Duolingo's Half-Life Regression model for spaced repetition, which estimates the memory half-life of individual word-user pairs and schedules review sessions to occur at the predicted moment of forgetting, significantly improving vocabulary retention compared to fixed-interval review schedules. A randomized evaluation of Duolingo's AI-optimized curriculum showed that 34 hours of practice on the platform produced Spanish proficiency gains equivalent to one semester of university-level instruction [26].

For English language learners (ELL) in mainstream classrooms, AI-powered speech recognition and pronunciation coaching systems provide formative feedback that would otherwise require individual teacher time impractical in large classes. Automatic speech recognition systems trained specifically on non-native speech have achieved word error rates below 10% for major language pairs [27], enabling reliable feedback on phonemic accuracy, prosody, and fluency that supports ELL students in developing oral proficiency alongside academic content knowledge.

6. AI-DRIVEN LEARNING ANALYTICS

A. Early Warning Systems and At-Risk Identification

Learning analytics refers to the measurement, collection, analysis, and reporting of data about learners and their contexts for the purpose of understanding and optimizing learning [28]. AI-powered learning analytics systems process multivariate streams of engagement data from learning management systems—including login frequency, content access patterns, assignment submission timing, discussion forum participation, and assessment performance—to build predictive models of student success and identify at-risk students early enough for timely intervention. Arnold and Pistilli [29] developed the Course Signals system at Purdue University, which used predictive modelling to identify students at academic risk and alert instructors to trigger targeted outreach, resulting in significantly improved retention rates and grade distributions compared to historical cohorts.

A systematic review by Hlosta et al. [30] of 40 early warning system implementations across higher education found that predictive models achieved AUC scores of 0.70–0.90 for end-of-course grade prediction when evaluated as early as the third week of a semester, providing substantially earlier identification of at-risk students than traditional midterm grade monitoring. However, the same review documented significant variation in intervention effectiveness, with institutions that paired early warning analytics with structured academic advising protocols achieving substantially better student outcomes than those that relied on algorithmic alerts without accompanying human support structures.

B. Teacher-Facing Analytics Dashboards

A critical success factor for learning analytics is the design of teacher-facing interfaces that translate complex model outputs into actionable instructional insights without overwhelming educators with data. Research on analytics dashboard design emphasizes the importance of making statistical uncertainty visible, presenting comparative reference points that enable teachers to contextualize individual student data within class and cohort norms, and aligning dashboard affordances with the specific decisions teachers make at different points in the instructional cycle [31]. Matcha et al. [32] conducted a mixed-methods study of teacher use of an analytics dashboard integrated into an LMS, finding that teachers who received structured professional development in data-informed instructional decision-making made significantly more adaptive instructional adjustments in response to analytics data than those provided dashboard access without training, underscoring that analytics tools require accompanying pedagogical support to be effective.

7. LARGE LANGUAGE MODELS IN EDUCATIONAL CONTEXTS

A. Generative AI as Instructional Tool

The release of GPT-3.5 and GPT-4 by OpenAI in 2022–2023 catalysed widespread experimentation with large language models as educational tools, sparking intense debate within educational research and policy communities regarding the implications for academic integrity, critical thinking development, and the future of pedagogical practice [33]. LLMs offer genuinely novel instructional capabilities including on-demand explanation of concepts at adjustable complexity levels, generation of practice problems and examples tailored to learner context, Socratic dialogue simulation that scaffolds higher-order thinking without revealing answers directly, and multi-lingual support that reduces language access barriers for diverse learners [34].

Empirical studies of LLM-assisted learning are accumulating rapidly. Bastani et al. [35] conducted a randomized controlled trial comparing students who used a GPT-4-based AI tutor with a control group receiving conventional instruction in a mathematics course, finding that AI-tutored students achieved learning gains 0.37 standard deviations larger than controls during the tutored phase. Notably, however, the study found that AI-tutored students who subsequently completed assessments without AI assistance showed smaller skill transfer gains than the control group, raising important questions about dependency formation and the conditions under which AI assistance promotes versus substitutes for deep learning.

B. Academic Integrity and Critical Thinking

The capacity of LLMs to generate high-quality academic text has precipitated a crisis in traditional essay-based assessment, undermining the validity of written assignments as measures of student learning in contexts where AI use cannot be reliably detected or controlled. AI-generated text detection tools including GPTZero and Turnitin's AI detection feature have demonstrated detection accuracies of 70–98% in controlled evaluations, but performance degrades substantially in real-world conditions with sophisticated prompting strategies, and false positive rates create risks of unjust academic misconduct allegations against legitimate student work [36]. The pedagogical response to this challenge has shifted from detection toward assessment redesign, emphasizing oral examinations, process portfolios, in-class writing, and multi-stage assignments that require authentic demonstration of learning not replicable through AI generation [37].

C. AI Literacy as Educational Imperative

The pervasive integration of AI into professional and civic life creates an imperative for educational systems to develop AI literacy—the conceptual understanding, critical evaluation skills, and practical competencies needed to use AI systems effectively and responsibly—as a core educational outcome alongside traditional literacies. Long and Magerko [38] proposed a framework for AI literacy comprising 17 competencies organized around understanding AI, using and applying AI, evaluating AI, and creating with AI. The integration of AI literacy into K-12 curricula requires significant teacher professional development investment, as surveys consistently indicate that the majority of current teachers feel insufficiently prepared to teach AI concepts or to critically evaluate AI tools for classroom adoption [39].

8. CHALLENGES AND ETHICAL CONSIDERATIONS

A. Equity, Access, and the Digital Divide

The benefits of AI-enhanced education are not equitably distributed. Access to high-quality AI educational tools requires reliable internet connectivity, up-to-date devices, and digital literacy competencies that remain unequally distributed across socioeconomic, geographic, and demographic lines. The COVID-19 pandemic starkly illuminated these inequities, as the rapid transition to technology-mediated instruction widened achievement gaps between students with and without reliable digital access [40]. AI educational systems trained predominantly on data from well-resourced educational contexts may embed implicit biases that produce suboptimal performance for learners from underrepresented linguistic, cultural, or socioeconomic backgrounds, a phenomenon documented in automated grading systems that show differential scoring accuracy across student demographic groups [41].

B. Data Privacy and Student Protection

AI educational systems are inherently data-intensive, requiring collection and processing of detailed behavioural data that may include sensitive information about student learning difficulties, affective states, and academic performance trajectories. The protection of this data is governed by regulations including the Family Educational Rights and Privacy Act (FERPA) in the United States, the Children's Online Privacy Protection Act (COPPA) for users under 13, and the GDPR in Europe [42]. Compliance complexity is heightened in educational contexts by the involvement of multiple stakeholders—schools, districts, platform vendors, and third-party analytics providers—creating governance gaps in which student data flows to parties beyond the awareness of students and parents. Trust frameworks for educational AI require clear data minimization principles, purpose limitation, meaningful consent mechanisms, and independent audit rights for educational institutions procuring AI tools.

C. Teacher Role and Professional Agency

A recurring concern in educational AI discourse is the risk of de-professionalizing teaching by over-automating instructional decisions that require the relational, contextual, and ethical judgment that constitutes teaching expertise. Research on teacher responses to AI tools documents a range of reactions from enthusiastic adoption to principled resistance, with teacher professional identity, perceived autonomy, and trust in system transparency as key moderating variables [43]. Pedagogical frameworks for AI integration in education emphasize the complementary relationship between AI and teacher capabilities—AI handling routine differentiation, data synthesis, and administrative tasks while teachers focus on mentorship, motivational support, complex problem facilitation, and the social and emotional dimensions of learning that AI systems cannot replicate [1].

9. FUTURE RESEARCH DIRECTIONS

Several high-priority research directions emerge from this review. First, longitudinal studies examining the long-term effects of AI-assisted learning on knowledge transfer, metacognitive skill development, and academic trajectory are critically needed, as the majority of existing evidence derives from short-term controlled experiments with limited ecological validity [8]. The question of whether AI-driven efficiency gains in the acquisition of declarative knowledge support or undermine the development of productive struggle, self-regulation, and deep conceptual understanding requires careful longitudinal investigation.

Second, the development of culturally responsive AI educational systems that perform equitably across diverse learner populations demands greater attention. Current ITS and adaptive learning platforms are predominantly developed and validated in high-income, English-language contexts, limiting their generalizability and raising concerns about the global exportation of culturally situated pedagogical assumptions embedded in AI systems [44]. Third, research on human-AI collaboration in teaching—examining how teachers integrate AI tools into their practice, what forms of professional development are most effective in supporting this integration, and how the teacher-AI collaboration relationship evolves over time—is essential for informing implementation policy [43].

Fourth, the governance of AI in education requires interdisciplinary research integrating educational policy, computer science, ethics, and law to develop robust frameworks for procurement, deployment, and audit of AI educational tools. The current landscape, characterized by rapid commercial innovation and lagging regulatory frameworks, creates risks for students and institutions that require proactive policy attention. Fifth, investigation of AI applications for supporting students with learning differences—including dyslexia, attention deficit disorders, autism spectrum conditions, and physical disabilities—represents a high-impact research frontier where AI personalization capabilities may be particularly valuable [45].

10. CONCLUSION

This review has examined the theoretical foundations, empirical evidence, and ethical dimensions of AI applications across the principal domains of educational practice. The evidence base is substantial and, on balance, supportive of carefully designed AI integration producing meaningful learning outcome improvements: intelligent tutoring systems achieve effect sizes comparable to one-on-one human tutoring; adaptive learning platforms improve both achievement and learning efficiency; automated feedback accelerates the revision process; learning analytics enable timely identification of and intervention for at-risk students; and large language models are opening new possibilities for personalized instructional dialogue at scale.

At the same time, this review underscores that AI is not a self-sufficient educational intervention. Its effectiveness is contingent on alignment with sound pedagogical theory, implementation fidelity, teacher professional development, equitable access infrastructure, and robust data governance. The most consequential risks—algorithmic bias, privacy violation, digital exclusion, and the erosion of teacher professional agency—are not inevitable features of educational AI but design and governance choices that policymakers, researchers, platform developers, and educational institutions must actively work to prevent.

The fundamental purpose of education—to support the intellectual, social, emotional, and ethical development of human beings—cannot be reduced to the optimization of measurable performance metrics that AI systems currently excel at maximizing. Realizing the transformative potential of AI in education requires keeping this broader purpose at the centre of design and implementation decisions, treating AI as a powerful tool in service of human flourishing rather than an end. As AI capabilities continue to advance, ongoing dialogue among technologists, educators, students, families, and policymakers will be essential to ensure that educational AI serves equity, quality, and human dignity alongside efficiency.

REFERENCES

[1] R. Luckin et al., Intelligence Unleashed: An Argument for AI in Education. London, U.K.: Pearson Education, 2016.

[2] J. R. Carbonell, "AI in CAI: An artificial-intelligence approach to computer-assisted instruction," IEEE Transactions on Man-Machine Systems, vol. 11, no. 4, pp. 190–202, Dec. 1970. doi: 10.1109/TMMS.1970.299942.

[3] W. Holmes, M. Bialik, and C. Fadel, Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Boston, MA, USA: Center for Curriculum Redesign, 2019.

[4] Global Market Insights, "AI in Education Market Size and Forecast, 2023–2032," Global Market Insights, Selbyville, DE, USA, Rep., 2023.

[5] B. P. Woolf, Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning. Burlington, MA, USA: Morgan Kaufmann, 2010.

[6] C. Piech et al., "Deep knowledge tracing," in Advances in Neural Information Processing Systems (NeurIPS), vol. 28, 2015, pp. 505–513.

[7] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," User Modeling and User-Adapted Interaction, vol. 4, no. 4, pp. 253–278, 1994. doi: 10.1007/BF01099821.

[8] K. VanLehn, "The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems," Educational Psychologist, vol. 46, no. 4, pp. 197–221, 2011. doi: 10.1080/00461520.2011.611369.

[9] B. S. Bloom, "The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring," Educational Researcher, vol. 13, no. 6, pp. 4–16, Jun./Jul. 1984. doi: 10.3102/0013189X013006004.

[10] J. A. Kulik and J. D. Fletcher, "Effectiveness of intelligent tutoring systems: A meta-analytic review," Review of Educational Research, vol. 86, no. 1, pp. 42–78, Mar. 2016. doi: 10.3102/0034654315581420.

[11] S. Ritter, J. R. Anderson, K. R. Koedinger, and A. Corbett, "Cognitive tutor: Applied research in mathematics education," Psychonomic Bulletin and Review, vol. 14, no. 2, pp. 249–255, Apr. 2007. doi: 10.3758/BF03194060.

[12] L. S. Vygotsky, Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA, USA: Harvard Univ. Press, 1978.

[13] C. Dziuban, C. R. Graham, P. D. Moskal, A. Norberg, and N. Sicilia, "Blended learning: The new normal and emerging technologies," International Journal of Educational Technology in Higher Education, vol. 15, no. 3, 2018. doi: 10.1186/s41239-017-0087-5.

[14] C. Tanes, K. E. Arnold, A. S. King, and M. A. Remnet, "Using Signals for appropriate feedback: Perceptions and practices," Computers and Education, vol. 57, no. 4, pp. 2414–2422, Dec. 2011. doi: 10.1016/j.compedu.2011.05.016.

[15] R. S. J. d. Baker, A. T. Corbett, K. R. Koedinger, and A. Z. Wagner, "Off-task behavior in the cognitive tutor classroom: When students game the system," in Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI), Vienna, Austria, Apr. 2004, pp. 383–390. doi: 10.1145/985692.985741.

[16] S. K. D'Mello and A. Graesser, "AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back," ACM Transactions on Interactive Intelligent Systems, vol. 2, no. 4, pp. 1–39, Jan. 2013. doi: 10.1145/2395123.2395128.

[17] E. B. Page, "The imminence of grading essays by computer," Phi Delta Kappan, vol. 47, no. 5, pp. 238–243, 1966.

[18] T. Ramesh and D. Sanampudi, "An automated essay scoring systems: A systematic literature review," Artificial Intelligence Review, vol. 55, no. 3, pp. 2495–2527, Mar. 2022. doi: 10.1007/s10462-021-10068-2.

[19] M. Shermis and J. Burstein, Eds., Handbook of Automated Essay Evaluation: Current Applications and New Directions. New York, NY, USA: Routledge, 2013.

[20] L. C. Perelman, "BABEL generator and e-rater: 21st century writing constructs and automated essay scoring," Journal of Writing Assessment, vol. 7, no. 1, 2014.

[21] M. Rankin, L. Abbing, and T. Lindner, "Automated formative feedback in writing: A quasi-experimental study of secondary school English," Computers and Education: Artificial Intelligence, vol. 4, p. 100110, 2023. doi: 10.1016/j.caeai.2022.100110.

[22] K. Crossley, S. Skalicky, and D. McNamara, "Moving beyond classic readability formulas: New methods and new models," Journal of Research in Reading, vol. 42, no. 3–4, pp. 468–485, 2019. doi: 10.1111/1467-9817.12283.

[23] X. Du, J. Shao, and C. Cardie, "Learning to ask: Neural question generation for reading comprehension," in Proc. 55th Annu. Meeting Assoc. Comput. Linguistics (ACL), Vancouver, BC, Canada, Jul. 2017, pp. 1342–1352. doi: 10.18653/v1/P17-1123.

[24] L. Shortt, Y. Tilak, I. Kuznetcova, B. Martens, and B. Akinkuolie, "Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020," Computer Assisted Language Learning, vol. 36, no. 3, pp. 517–554, 2023. doi: 10.1080/09588221.2021.1933540.

[25] B. Settles and B. Meeder, "A trainable spaced repetition model for language learning," in Proc. 54th Annu. Meeting Assoc. Comput. Linguistics (ACL), Berlin, Germany, Aug. 2016, pp. 1848–1858. doi: 10.18653/v1/P16-1174.

[26] P. Vesselinov and J. Grego, "Duolingo effectiveness study," City Univ. of New York, New York, NY, USA, Rep., 2012.

[27] A. Hinsvark, L. Jalbert, and J. Hansen, "Non-native speech recognition for language learning: A survey," in Proc. 17th Int. Conf. Spoken Language Processing (Interspeech), San Francisco, CA, USA, Sep. 2016, pp. 2624–2628. doi: 10.21437/Interspeech.2016-474.

[28] Society for Learning Analytics Research (SoLAR), "1st International Conference on Learning Analytics and Knowledge," Banff, AB, Canada, 2011.

[29] K. E. Arnold and M. D. Pistilli, "Course signals at Purdue: Using learning analytics to increase student success," in Proc. 2nd Int. Conf. Learning Analytics and Knowledge (LAK), Vancouver, BC, Canada, 2012, pp. 267–270. doi: 10.1145/2330601.2330666.

[30] M. Hlosta, L. Zdrahal, and J. Zendulka, "Ouroboros: Early identification of at-risk students without models based on legacy data," in Proc. 7th Int. Learning Analytics and Knowledge Conf. (LAK), Vancouver, BC, Canada, Mar. 2017, pp. 6–15. doi: 10.1145/3027385.3027449.

[31] S. Schwendimann et al., "Perceiving learning at a glance: A systematic literature review of learning dashboard research," IEEE Transactions on Learning Technologies, vol. 10, no. 1, pp. 30–41, Jan.–Mar. 2017. doi: 10.1109/TLT.2016.2599299.

[32] W. Matcha, D. Gasevic, N. A. Uzir, J. Jovanovic, and A. Pardo, "Analytics of learning strategies: Associations with academic performance and feedback," in Proc. 9th Int. Learning Analytics and Knowledge Conf. (LAK), Tempe, AZ, USA, Mar. 2019, pp. 461–470. doi: 10.1145/3303772.3303787.

[33] M. Baidoo-Anu and L. O. Ansah, "Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning," Journal of AI, vol. 7, no. 1, pp. 52–62, 2023. doi: 10.61969/jai.1337500.

[34] D. Mollick and L. Mollick, "Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts," SSRN Working Paper, Mar. 2023. doi: 10.2139/ssrn.4391243.

[35] H. Bastani et al., "Generative AI can harm learning," Working paper, The Wharton School, Univ. of Pennsylvania, Philadelphia, PA, USA, 2024.

[36] L. Gao et al., "Human-like summarization evaluation with ChatGPT," arXiv:2304.02554, Apr. 2023.

[37] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education," International Journal of Educational Technology in Higher Education, vol. 20, no. 43, 2023. doi: 10.1186/s41239-023-00411-8.

[38] D. Long and B. Magerko, "What is AI literacy? Competencies and design considerations," in Proc. CHI Conf. Human Factors in Computing Systems (CHI), Honolulu, HI, USA, Apr. 2020, pp. 1–16. doi: 10.1145/3313831.3376727.

[39] S. Celik, "Toward a unified definition of artificial intelligence literacy for educational settings," Computers and Education: Artificial Intelligence, vol. 4, p. 100082, 2023. doi: 10.1016/j.caeai.2022.100082.

[40] UNICEF, "COVID-19 and school closures: One year of education disruption," UNICEF, New York, NY, USA, Rep., Mar. 2021.

[41] B. Bridgeman, C. Trapani, and Y. Attali, "Comparison of human and machine scoring of essays: Differences by gender, ethnicity, and country," Applied Measurement in Education, vol. 25, no. 1, pp. 27–40, 2012. doi: 10.1080/08957347.2012.635502.

[42] U.S. Department of Education, "Family Educational Rights and Privacy Act (FERPA)," 20 U.S.C. § 1232g; 34 CFR Part 99, Washington, DC, USA, 1974 (as amended).

[43] I. Selwyn, Should Robots Replace Teachers? AI and the Future of Education. Cambridge, U.K.: Polity Press, 2019.

[44] J. Lim, L. Lim, and B. Lim, "Artificial intelligence in education and its culturally relevant considerations," Journal of Research in Innovative Teaching and Learning, vol. 16, no. 1, pp. 54–68, 2023. doi: 10.1108/JRIT-06-2022-0034.

[45] N. Drigas and P. Ioannidou, "Artificial intelligence in special education: A decade review," International Journal of Engineering Education, vol. 28, no. 6, pp. 1366–1372, 2012.