

The Role of Generative Artificial Intelligence in Personalised Learning: Opportunities, Challenges, and Future Directions in Higher Education

Kerong Huang

Faculty of Education, The University of Hong Kong, Hong Kong, China

Abstract: Generative artificial intelligence (GenAI), especially large language models, has enhanced personalised learning by enabling dynamic content creation, conversational assistance, adaptive feedback, and fast resource creation. Nevertheless, technical capability cannot be translated into educational value. This systematic literature review explores the potential of GenAI to improve personalised learning in higher education and determines technological, pedagogical, ethical, and institutional circumstances of implementation. A systematic search of Web of Science and PsycINFO yielded 201 records; following the elimination of duplicates and sequential screening, 18 articles published in 2020-2025 were selected. Thematic synthesis and co-occurrence mapping revealed opportunities in learner-specific content, adaptive tutoring, self-directed learning, accessibility, assessment support, and educator productivity. Constant threats involved hallucinated or biased results, privacy and copyright issues, academic-integrity pressures, unequal access, and cognitive offloading that can undermine critical engagement. There is a lot of evidence that is conceptual, short-term, or grounded on perceived usefulness as opposed to long term learning outcomes. The review supports a human-centred paradigm in which GenAI does not replace but complements teachers. To adopt it responsibly, one needs verified content, redesigning of assessments, open governance, fair access, and clear AI literacy among staff and students. These factors define whether GenAI can offer valuable scaffolding or simply produce answers.

Keywords: Generative artificial intelligence, personalised learning, adaptive learning, higher education, educational technology, AI literacy.

1. Introduction

Digital technologies have reshaped higher education by changing how knowledge is produced, accessed, assessed, and communicated. Artificial intelligence (AI) has contributed to this shift through automated assessment, administrative support, and learning analytics [1]. The more recent development of Generative Artificial Intelligence (GenAI) extends these functions. Instead of only classifying information or predicting outcomes, GenAI systems can produce text, images, audio, code, and other outputs in response to prompts by modelling patterns in large datasets [2,3]. This generative capacity has intensified interest in whether AI can make learning more responsive to individual students.

Personalised learning begins from the premise that learners differ in prior knowledge, pace, goals, preferences, and support needs [4]. In principle, teaching should therefore adapt content, sequencing, feedback, and challenge to the learner. In practice, sustained personalisation is resource intensive. Instructors must diagnose needs, develop differentiated resources, respond promptly, and monitor progress, often for large and diverse cohorts [5, 6]. Earlier adaptive systems partly addressed this problem through pre-programmed pathways, whereas GenAI can generate context-sensitive explanations, examples, practice tasks, and dialogue in real time.

Current applications suggest several possible benefits. Students can request simplified or alternative explanations, generate practice questions, rehearse professional scenarios, and receive immediate feedback [7]. GenAI can assist educators in drafting learning content, generating formative assessment questions, and finding methods to differentiate

instruction [8, 9]. AI-enhanced environments can also facilitate self-directed learning by assisting students in setting goals, finding resources, and tracking their progress [10]. These are relevant in the context of higher education where independent learning is required but individual tuition is not always available.

However, personalised production is not the same as educationally efficient personalisation. GenAI can generate fluent but inaccurate output, which is often referred to as hallucination [11]. Its responses can reproduce biases embedded in training data, while the processing of learner information raises privacy, transparency, and accountability questions [12]. The fact that it can produce polished academic writing also breaks the traditional assumptions regarding authorship and evaluation. On the pedagogical level, the ability to generate answers quickly can decrease productive struggle and make students outsource interpretation, evaluation, and writing instead of building these skills themselves [12].

There is a tension in the literature. GenAI increases access to feedback and explanations, but can mask gaps in understanding with seemingly reasonable responses. It decreases repetitive work, but introduces new work in verification, governance, assessment design and support for learners. It may increase equity through translation and accessibility, but high-quality tools, infrastructure and varying levels of AI literacy may exacerbate inequities [13, 14]. The impact of GenAI on personalised learning hinges on how technological affordances, pedagogic approaches, institutional protections, and learner autonomy interact.

This tension is important as personalised learning tends to be overused. Individualisation of content occurs when the wording, pace or examples vary; individualisation of learning

occurs only when changes are made that respond in a meaningful way to the evolving understanding of learners. Systems which produce unlimited variants can add choice without enhancing diagnosis or progression. Assessment should then take into account alignment with curricula, quality of adaptation and independent performance following withdrawal of support [15].

The practices of higher education are varied. The usefulness and dangers of GenAI differ in terms of exploratory brainstorming, mathematical reasoning, clinical judgement, coding, language learning, and complex academic writing. These differences are ignored by blanket institutional claims. Therefore, the purpose of the task, stakes, and human oversight are used to interpret the reviewed literature rather than as an argument in favor of unrestricted adoption.

There has been a rapid increase in research in this field, yet evidence is scattered between tutoring, feedback, content creation, assessment, professional education, and ethical commentary. The current reviews define the applicability of GenAI to cognitive schema construction and adaptive systems [6, 15], yet a more specific synthesis is required to relate opportunities to implementation conditions and perceived usefulness versus demonstrated learning. This review addresses this need by answering the following question: How does GenAI improve personalised learning in higher education, and what are the opportunities and challenges that impact its successful application?

The review aims to achieve two goals: to map and synthesise the opportunities and challenges of GenAI-driven personalised learning, and to provide recommendations for responsible integration in higher education. Its thesis is that GenAI is a complementary technology. It complements feedback, scaffolding and resource development, but does not replace teachers, who continue to make pedagogic decisions, verify information, provide relational support, and make ethical judgement.

2. Literature Review

2.1. GenAI as an Enabler of Personalised Learning

Personalised learning seeks to adjust educational experiences to learner characteristics rather than deliver identical instruction to all students [4]. Before GenAI, adaptive platforms generally selected from materials and pathways designed in advance. GenAI changes the scale and speed of adaptation because it can create new explanations, examples, summaries, questions, simulations, and multimedia resources on demand [6, 8]. This capacity may reduce the practical burden of producing differentiated materials for large cohorts.

The most persuasive use case is not unrestricted content generation, but targeted variation within a teacher-designed curriculum. For example, a learner may ask for a concept to be explained through a worked example, an analogy, or a discipline-specific scenario. Studies of customised learning paths and tutoring personas suggest that GenAI can adjust sequencing, tone, and difficulty to learner needs [16]. In medical and professional education, scenario generation may support rehearsal of decision-making that would otherwise be costly or difficult to reproduce [17].

GenAI may also broaden the formats through which students encounter a concept. Text can be converted into structured summaries, questions, dialogues, or simulated

cases; complex wording can be simplified; and examples can be contextualised to a student's programme or interests [6, 8, 14]. Such variety can support access and comprehension, particularly when a single explanation has failed. The benefit is strongest when multiple representations lead back to shared disciplinary standards rather than fragmenting the curriculum into disconnected, individually generated content.

From an institutional perspective, scalable generation appears to address the cost of differentiation. Yet scale introduces a quality-control problem. A lecturer can review a fixed resource before release, whereas learner-specific output is produced at the moment of use. The burden of checking shifts toward the learner unless systems are constrained by verified sources or monitored by staff. Personalisation at scale therefore requires stronger assurance processes, not fewer.

However, personalisation can be superficial. A response may sound tailored because it mirrors a prompt while remaining poorly aligned with course outcomes or the learner's misconceptions. Systems infer needs from prompts and available data; they do not possess a complete understanding of the learner. Evidence is also uneven. Much of the literature describes possible applications or reports favourable perceptions, while relatively few studies measure retained knowledge, transfer, or performance over time [12, 18]. GenAI should therefore be treated as an infrastructure for personalisation rather than as an autonomous provider of personalised education.

2.2. Adaptive Feedback, Tutoring, and Learner Agency

Feedback is central to learning when it clarifies current performance, desired performance, and the actions required to close the gap [19]. GenAI has the potential to make feedback more real-time and conversational, enabling learners to pose follow-up questions and ask to see additional examples. The studies of generative chatbots, adaptive coding assistance, virtual tutors, and formative assessment reveal speed, availability, and flexible explanation as common strengths [9, 20]. These features may extend support beyond scheduled contact hours and help learners persist when an instructor is unavailable.

The quality of feedback still relies on accuracy, specificity and alignment with criteria. A smooth answer may still be wrong in diagnosing an error or rewarding a feeble strategy. Automated feedback should be validated, limited by definite learning outcomes, and revised in time instead of giving a complete solution. This difference is essential when it comes to assessment, where GenAI can personalise questions and feedback but makes it difficult to isolate actual performance and assisted production [21, 22].

Within feedback research, there is a difference between evaluative comments and actionable guidance. Even a fluent response may misdiagnose an error or reward a weak strategy. Automated feedback should be validated, limited by clear learning outcomes and revised in time instead of giving a completed answer. This difference is essential in assessment, as GenAI tailors questions and feedback but makes it difficult to distinguish between real performance and assisted production [9, 19].

Self-directed learning (SDL) can also be facilitated with the help of GenAI. It can be used by students to set goals, plan studies, create questions, alternatives and self-reflection [6, 10, 23]. Such activities are able to reinforce agency in case learners make decisions and assess output. On the other hand,

agency can decrease when students do not critically accept generated responses. The same system may work as a thinking partner or replacement. Pedagogical design should then require that students justify decisions, contrast sources, confirm assertions, and record the utilization of AI.

Agency is therefore a design decision. GenAI can enhance control over learning when learners make decisions about what to ask, evaluate competing responses, and update their knowledge. When the system suggests goals, chooses resources, generates answers, and assesses work without transparency, the superficial personalisation may turn into algorithmic guidance. Recommendations should be contestable through interfaces and teaching practices that maintain meaningful learner choice.

It is also important that educators are supported. GenAI can speed up the creation of routine resources, propose differentiated tasks, and help with formative assessment [8, 24]. However, efficiency arguments must be qualified: teachers need to learn tools, check outputs, safeguard data, redesign assessment, and address new types of misconduct. The time saved in writing can be allocated to verification and governance. Collaborative model is thus the best, where automation assists repetitive or generative work whereas educators offer disciplinary standards, context and relational judgement [25].

2.3. Challenges and Limits of Integration

The most immediate technical issue is reliability. GenAI systems have the potential to produce inaccurate, fabricated or internally inconsistent information and present it confidently [11, 12]. With personalised environment, errors might not be easily noticed since every learner will get varied output. Verification should not be a supplement; institutions require sanctioned use cases, source-checking processes, and escalation to human expertise, especially in high-stakes or professionally regulated topics [17, 26].

Claims of inclusive personalisation are complicated by bias and privacy. Stereotypes or lack of representation of specific cultures, languages, and ways of knowing can be reproduced in training data. Personalisation might also necessitate sensitive data regarding performance, disability or behaviour. In the absence of open data policies, informed consent and responsible procurement, systems may generate discriminatory suggestions or reveal learner information [14,21]. The accessibility functions like translation, text simplification, and multimodal output are still useful, yet their usefulness is determined by quality and fair access to effective tools [13, 14].

The concept of academic integrity is often discussed in terms of detection, but GenAI reveals a problem with assessment design. When a task can be performed using generic text generation, it might not be a good indicator of individual reasoning. Other more justifiable methods are staged submissions, oral explanation, authentic application, process records and reflective commentary on AI use [9,21]. These strategies do not remove misuse, but bring assessment closer to learning processes.

Copyright and Authorship are also problems. Generated material may indirectly use protected training data, and it is unclear whether the user owns it or how to attribute and use it. Be open about using artificial intelligence to help write this work and take responsibility for verifying the sources and facts. Institutional guidance should guide practice in light of changes in tools and laws [14,21].

Lastly, adoption is subject to institutional preparedness. Teachers and learners require AI literacy, such as the capability to prompt, verify, cite, disclose, and challenge generated content. The costs, disparity in infrastructure, accessibility, workload, and disciplinary differences must also be addressed by institutions [24, 27]. Research on Technology Acceptance Model indicates that perceived usefulness and ease of use are factors that affect adoption [28], yet educational institutions should not focus on acceptance only: a tool may be easy to use and lack pedagogical strength or be ethically unacceptable.

3. Theoretical Framework

The analysis is based on four complementary approaches: constructivism and cognitive schema theory, Generative Learning Theory (GLT), the Zone of Proximal Development (ZPD) and the Technology Acceptance Model (TAM). They combine learning processes with implementation conditions together.

Constructivism is a theory that believes that learners actively construct knowledge as they interact with information, experience and others [29]. This knowledge is organised in cognitive schemas which evolve as learners incorporate new information. GenAI can aid schema development by generating examples, representations, and questions that relate the unknown material to what is already known [6]. It can sabotage the same process when students are fed generated explanations without trying, organising or using them.

GLT suggests that meaningful learning involves the selection, organisation and integration of relevant information with prior knowledge [30]. This is facilitated by GenAI through prompts for explanation, comparison, summarisation and questioning. This explains why answer substitution is problematic: the model's output does not guarantee the learner did the thinking.

ZPD describes the distance between what a learner can do independently and what can be achieved with appropriate support [31]. Conversational tutors can provide hints, questions, and graduated explanations, making GenAI a potential source of scalable scaffolding. However, effective scaffolding is contingent and gradually withdrawn; current systems may over-assist, misunderstand emotional cues, or fail to recognise when a learner needs human intervention.

All these learning theories focus on one criterion: learners need to be intellectually active. Constructivism is about constructing knowledge, GLT generating processing and ZPD calibrated assistance. Both predict GenAI features are beneficial or detrimental based on task design. A generated analogy will trigger prior knowledge, but a final answer will not require organising knowledge.

TAM describes adoption in terms of perceived usefulness and ease of use [28]. It explains why students and staff may adopt accessible conversational technologies, but it is not sufficient to ensure quality in education. In this review, TAM is therefore applied alongside learning theory: we need adoption for impact, but pedagogic fit, trustworthiness and governance for positive impact.

4. Method

4.1. Review Design and Eligibility

A systematic literature review (SLR) was conducted to find and synthesise evidence on GenAI and personalised learning.

SLRs offer a clear approach to searching, selecting, appraising and synthesising evidence [32 - 34]. The use of GenAI in education is an emerging area, so empirical, conceptual, perspective and review studies were included in the review. This inclusive eligibility criterion enabled the synthesis to include early evidence of implementation and development of theory, but study design was taken into account when evaluating claims.

The studies that were eligible discussed GenAI in an educational setting, personalised or adaptive learning, and were published in English within 2020-2025 and aimed at higher education. Researches that concentrated on technical development alone without any educational use or those with no substantial relationship to personalisation were not considered.

The published study was the unit of analysis, rather than a particular tool. This was indicative of fast product changes that allowed comparing educational functions (feedback, tutoring and pathway design) across systems. In cases where papers integrated higher education with broader contexts, only results pertaining to post-secondary learning were synthesised.

4.2. Search Strategy and Selection

The search was conducted in Web of Science and

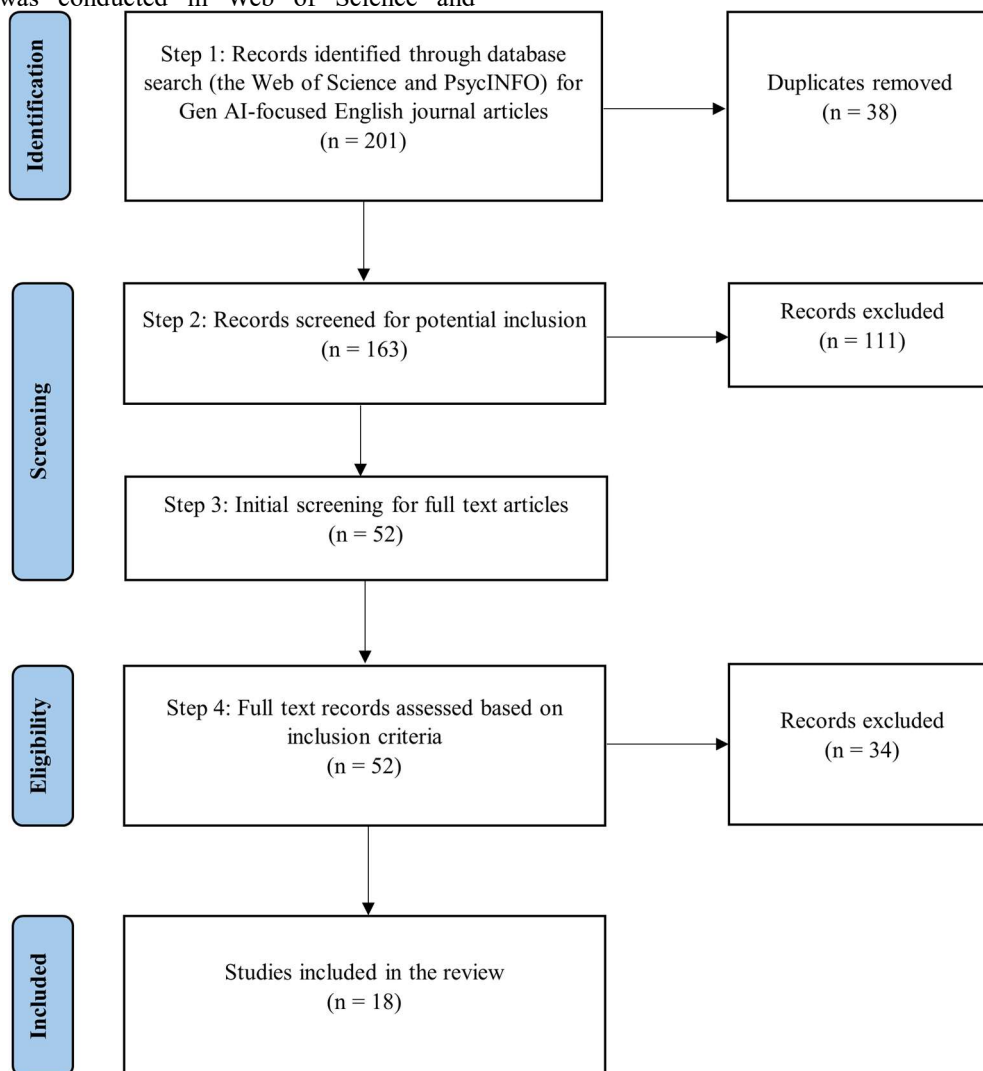


Figure 1. Systematic review search and selection process.

4.3. Quality Appraisal and Synthesis

Quality appraisal was adapted to the heterogeneous evidence base. Empirical studies were considered in relation

PsycINFO since their coverage areas are combined into education, psychology, technology, and interdisciplinary research. The plan involved three groups of concepts using Boolean operators:

GenAI: “generative AI”, “generative artificial intelligence”, “GPT”, “large language model”, or “LLM”;

Personalised Learning: “personalized learning”, “personalised learning”, “adaptive learning”, or “individualised instruction”; and

Education: “education”, “teaching”, “learning”, “classroom”, or “lesson”.

Selection followed the PRISMA sequence of identification, screening, eligibility assessment, and inclusion [35]. The search returned 201 records. After 36 duplicates were removed, 165 titles and abstracts were screened; 31 full texts were assessed against the eligibility criteria, and 18 studies were retained. Figure 1 summarises the process.

Screening decisions were based on explicit relevance to both GenAI and personalisation. Papers that discussed AI in education only in general terms were not retained unless they addressed adaptation, learner-specific support, or differentiated educational processes. This reduced the risk that the synthesis would equate any educational use of ChatGPT with personalised learning.

to clarity of aims, sampling, data collection, analytical appropriateness, and acknowledged bias, drawing on Critical Appraisal Skills Programme principles. Conceptual and

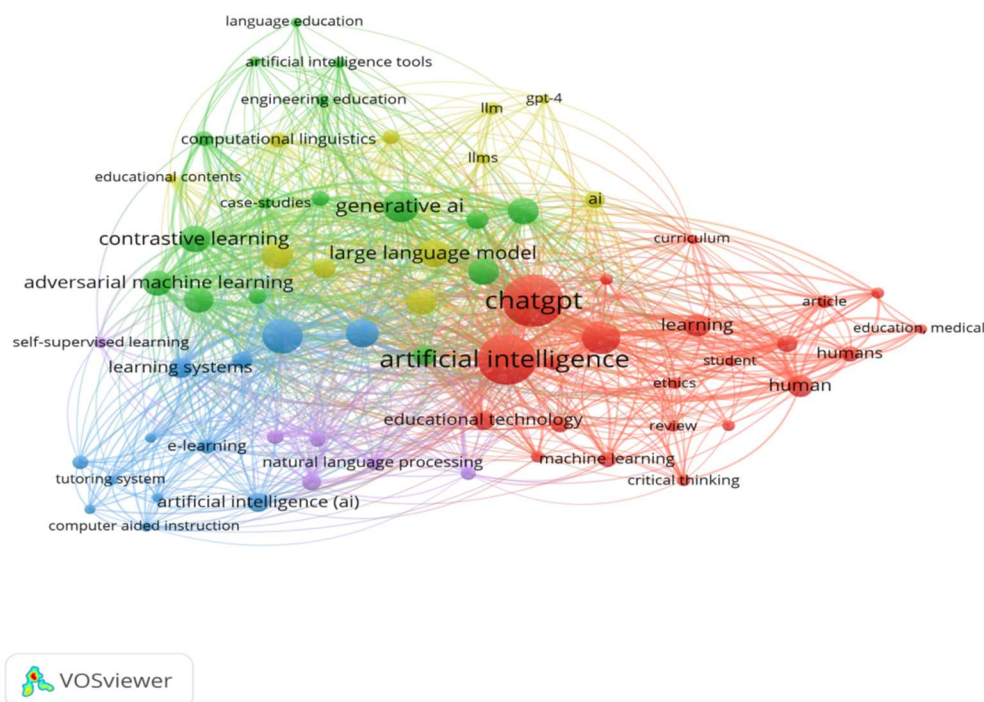
perspective papers were assessed for transparency of argument, evidential support, and relevance to the review question. No single numerical score was used across incompatible designs.

Thematic synthesis followed Thomas and Harden's staged approach of coding findings, developing descriptive themes, and constructing analytical themes [36]. Evidence was organised around opportunities and challenges, while attention was paid to study design and outcome type. VOSviewer was also used to examine co-occurrence patterns in the literature[37]. The mapping was treated as descriptive context rather than as evidence of educational effectiveness.

5.1. Characteristics and Thematic Structure of the Evidence

survey research [10,24], small experimental or mixed-method studies [38 - 39], case-based work on customised pathways[23], and analysis of changing teacher roles [25]. Contexts ranged from general higher education to supply-chain management, coding, pathology, microbiology, nephrology, and foreign-language education.

Empirical research differed in success measures, investigating perceptions, engagement and adoption, as well as feedback, performance and self-directed learning. These are different outcomes. Students may benefit from an AI tutor without learning more; teachers may save time writing but spend more time checking. Outcomes thus decouple affordances from effects.



5.2. Opportunities for Personalised Learning

of learner-specific resources, curricula, examples, and tutoring personas, together with real-time adjustment of difficulty or sequence [15, 17, 38]. This flexibility may be particularly useful where students enter a course with varied

prior knowledge. The evidence nevertheless focused more often on feasibility and perceived utility than on whether algorithmically customised materials improved retention or transfer.

A second theme was feedback and tutoring. GenAI-supported tools offered immediate explanations, hints, questions, and formative feedback [21,27]. In the stronger examples, students interacted iteratively with the system rather than receiving a single answer. This capability can increase the frequency of low-stakes practice and make support available outside teaching hours. The literature also cautioned that speed and fluency are not proxies for accuracy; feedback quality still requires criteria, verification, and opportunities for learner response[12,19].

GenAI was also associated with engagement and SDL as supported by the evidence. Practice can be made more interactive through conversational agents and virtual tutors, and exploration based on prompts can assist students in planning, questioning, and tracking learning[10,38]. Nevertheless, positive engagement was frequently evaluated based on perceptions or immediate use. The review thus differentiates behavioural interaction with a tool and deeper cognitive interaction with disciplinary concepts.

In studies, the prompting skill of the learner influenced the quality of interaction. Particular cues such as purpose, level and constraints elicited more usable support as compared to vague requests. This implies that results are partially based on previous digital confidence as opposed to system capability per se. Students who have better academic and technical skills benefit most without prompting and verification instruction[14].

Translation, language simplification, alternative formats, image descriptions, and scalable access to explanations were discussed as accessibility and inclusion[14, 21]. These functions can be advantageous to multilingual learners and students who need various representations of content. However, accessibility assertions were followed by the issue of unequal devices, connectivity, paid levels of models, and AI literacy. GenAI has the potential to reduce certain barriers and introduce others.

Lastly, GenAI can assist teachers with content creation, curriculum design, formative evaluation, and administrative support [17,25]. The advantage is the most likely to occur in first drafts and routine variation where professional review is still possible. The researches failed to justify taking educators out of the process but rather they explained a transition to curation, verification, coaching and ethical control.

These opportunity themes are interdependent. Content generation provides flexible resources; tutoring and feedback directs interaction; SDL enhances learner control; accessibility expands participation; educator support scales provision. A single weak aspect compromises others. Fast creation is not worth it when the feedback is inaccurate, not evaluated or not verified by educators.

5.3. Challenges Affecting Implementation

The most common concern was shared technical reliability. The hallucinated facts, unstable responses, and training data bias can deceive learners and destroy the trust [6,21,38]. Personalisation increases this difficulty since the output will differ among learners and might not be checked prior to utilization. Unrestricted use is particularly hard to justify in fields where mistakes are professionally consequential[26].

Privacy, copyright, transparency, accountability and

academic integrity were some of the ethical concerns [13,16,39]. The literature did not provide one solution to governance, but it was always pointing out that institutions should have clear policies regarding acceptable use, disclosure, data handling and human responsibility. Detection-only solutions were considered inadequate since they fail to consider why existing assessment tasks are susceptible to generated output.

Pedagogical issues revolved around cognitive offloading and over-dependence. Students can get polished answers without engaging in reasoning, writing or problem solving and this gives the false impression of performance that cannot be supported by knowledge[14,38]. Another concern was reduced human interaction especially when tutoring is understood as merely delivery of information but not a relational and motivational practice[25].

Equity and adoption challenges cut across the evidence. Access to capable tools, reliable infrastructure, and training is uneven[10,13]. Educators reported or were expected to need substantial professional development, while institutions faced costs of integration, maintenance, evaluation, and governance [26-27]. These findings suggest that implementation capacity, not technological availability alone, will shape who benefits.

Problems are systemic. Technical issues become pedagogical issues if students aren't skilled at detecting errors; unequal access becomes assessment issues when AI support is expected; ambiguous policies create workloads for instructors. Avoiding technical solutions, implementation needs to consider systems, teaching, assessment, ethics and support.

6. Discussion

6.1. Interpreting the Educational Value of GenAI

Results indicate that GenAI enhances the range, responsiveness and diversity of personalised learning but does not deterministically enhance learning. Its value is contingent: GenAI quickly produces alternatives that need to be assessed by educators and learners in terms of accuracy, relevance, challenge, and alignment. Personalisation is consequently a pedagogical relationship facilitated by technology, rather than an attribute of generated content.

This difference reconciles contradictory assertions in the literature. Close feedback breaks down barriers and facilitates persistence particularly when there is a lack of human tutoring. But convenience diminishes productive effort when systems do the intellectual work. It is not merely the use of GenAI, but the remaining cognitive work. Designs that involve learners comparing outputs, justifying decisions, finding evidence, revising drafts or describing errors are more appropriate to learning than those that reward unconsidered output.

The findings complicate workload stories. GenAI simplifies the creation of examples, quizzes, cases and initial feedback. Reliable deployment still requires review, policy formation, assessment redesign and student support. Some workload is decreased, some is changed. The institutions need to evaluate workload during the entire implementation cycle rather than use generation speed as an indicator of efficiency.

Another implication relates to the nature of expertise in education. When information production is cheap, the unique contribution of educators becomes choosing valuable problems, ordering challenge, diagnosing errors, maintaining motivation and determining when exceptions should be made.

These are not ancillary activities; they are key to personalisation. Thus, GenAI might enhance the importance of pedagogic judgement while automating aspects of content creation [24].

The review also recommends a careful move towards democratisation. Conversational support at low cost can help students who have not been previously tutored, but the effectiveness depends on the language coverage and accessibility, the quality of the device, and the capability to differentiate between reliable and unreliable results. Although interventions are capable of increasing average access, they can increase disparities between students who use the tool strategically and those who do not [13-14]. Equity is therefore not to be assumed but must be empirically studied.

6.2. Theoretical Implications

Constructivism and schema theory indicate why varied examples and dialogue can be useful: they may help learners connect new information with prior knowledge and reorganise misconceptions [6,29]. Yet generated explanations contribute only when learners actively interpret and apply them. A system that provides finished answers may appear personalised while bypassing knowledge construction.

GLT explains this process. The learner activities are selection, organisation and integration [30]. These processes can be triggered with GenAI to produce contrasting cases, questions or incomplete explanations that learners can evaluate. Its educational worth is reduced when it passively performs selection and synthesis to the students. Based on this, teachers should create AI-assisted activities based on transformation and verification rather than reproduction.

ZPD positions GenAI as potential scaffolding [31]. Immediate feedback, adaptive explanations and conversational practice can help with tasks just beyond independent level. Good scaffolding is dynamic and transitory. GenAI might be too helpful, not recognise confusion or persist when a human tutor would stop. Humans are needed to fine-tune challenge and determine the time to withdraw support.

TAM explains why conversational systems can be adopted quickly: they are accessible and visibly useful [28]. The results reveal the boundaries of acceptance as a success criterion. The high uptake can be accompanied by unreliable information, poor learning design or unfair access. Assessment must then integrate user acceptance and learning outcomes, ethical protection and distributional impacts.

6.3. Strength and Limitations of the Evidence

The evidence base is encouraging and methodologically immature. First, the corpus is filled with new publications, conceptual analyses, and context-related demonstrations. There were not many studies with comparative designs, and experimental samples could be small [18]. Second, perceived usefulness, satisfaction, and intention to use were more prevalent than objective indicators of retained knowledge, transfer or independent performance. Learning is not evidenced by positive attitudes.

Third, interventions were short-term. The long-term implications on critical thinking, writing development, dependence of the learner and workload of the teacher are unknown. The speed of model changes makes replication even more difficult: a study can assess a version, interface or policy that will change shortly. To enable the interpretation and repetition of findings, researchers should record versions

of models, prompts, system settings, conditions of comparison, and human support.

Improvement is also needed in outcome measurement. The quality of an AI-assisted product should be isolated in future research, as well as what a learner can subsequently do without assistance. Tests that are delayed, transfer tasks, justification of reasoning, and independent performance would be more convincing than the immediate accomplishment of the task. Process data can help to understand whether students interrogate ideas or avoid them with the help of GenAI.

Fourth, the use of conceptual and review literature was necessary given the infancy of the field but reduces causal inferences. The synthesis pinpoints mechanisms and conditions for implementation, not a specific effect size. Publication bias is also likely as innovative or positive uses may be more likely to be reported. This suggests caution: the evidence supports potential and conditional benefit, not guaranteed transformation.

6.4. A Human-Centred Implementation Framework

The pedagogical purpose, bounded technology, human oversight, and ongoing evaluation should be integrated into a responsible implementation model. To start with, institutions must start with a specific learning problem as opposed to an overall directive to adopt AI. Appropriate examples can involve low-stakes practice, alternative explanations, formative feedback drafts, language support or scenario generation. There is a need to have more robust controls in high stakes decisions and professionally sensitive guidance.

Second, teachers are to define what the system can produce, what students need to perform and how the output will be verified. Limitations, verification, attribution, disclosure, bias, privacy, and prompt design should be included in AI literacy. Students are supposed to realize that fluent output is temporary and accountability of work submitted is human.

Third, evaluation should bring thought into the open. Staged drafting, source trails, oral defence, authentic application, reflective commentary and comparing AI outputs are more convincing than one polished product. The policies must differentiate between acceptable help and substitution and be realistic in their interdisciplinary nature.

Fourth, institutions need to track outcomes and equity. Learning, workload, accessibility, error rates, dependence and access should be evaluated, rather than adoption. Data protection, transparency, accountability and challenge mechanisms are needed in procurement and governance. The outcome is a hybrid solution: GenAI will increase support, but educators will still have to ensure coherence, standards, relationships, and judgement.

Implementation ought to be progressive. Small pilots that have specific groups expose errors, workload, barriers and unintended behaviors early. Revision or withdrawal should be instigated by feedback in case the benefits are not worth the costs. This does not view innovation as a unidirectional shift towards automation but rather as an evaluative process.

There must be clear boundaries in the framework. Confidential data, clinical or legal judgement, or automated grading uses that are high stakes require more rigorous scrutiny than low-stakes brainstorming. Clear limits make it clear that responsible adoption is risk-based and not rule-based.

7. Limitations and Future Research

There are limitations to this review. Web of Science and PsycINFO searches offered interdisciplinary searching but may have missed studies indexed in other databases. The inclusion of only English-language publications may have limited the geographic and cultural diversity of the evidence. GenAI is rapidly evolving, so tools and findings may be quickly outdated. Lastly, study designs were not consistent, precluding meta-analysis; narrative synthesis was used.

Future studies should focus on longitudinal and comparative designs with objective measures of learning, transfer and independent performance. Researchers should provide details about model versions, prompts, human support and assessment conditions for replication. More emphasis is required on disciplinary differences, students with disabilities, multilingual education, low-resource institutions and the distribution of benefits and harms. Studies should also investigate AI-literacy interventions and compare human-AI configurations, rather than simply asking if GenAI is useful.

Language was also a limitation to the review process in a fast changing industry. The terms GenAI, adaptive learning, personalised learning and intelligent tutoring are not synonyms but overlap. Other studies can refer to it using other terms and some articles refer to generic content generation as personalisation. Search vocabulary should be updated in future reviews and more acute conceptual distinctions should be made as the field matures.

This evidence base would be enhanced by meta-scientific practices. Selective reporting would be minimized by preregistered hypotheses and analysis plans, shared prompts and materials, open reporting of failed implementations, and cross-institutional replication to enhance comparability. Multi-site studies are particularly useful because the implementation is highly dependent on institutional resources and student populations.

8. Implications for Higher Education

For teachers, results indicate that selective integration is indicated. GenAI can be defended when it comes to the expansion of practice and feedback without eliminating the necessity of students to reason, verify, and create. Course teams need to determine where human interaction is needed and where automated support would be beneficial.

Adoption needs to be a coordinated policy for institutions, not experimentation. Staff development, guidance, accessibility, procurement, privacy, assessment and quality assurance must come together. Support should be equitable, and personalised learning should not rely on a student's ability to afford a model that can perform it.

GenAI is an imperfect learner partner. Effective use involves purposeful questioning, verifying sources, identifying bias, and disclosing assistance. This makes AI literacy part of academic literacy, not just a technical skill.

The review suggests assessment leaders shift from policing tools to validating learning evidence. Assessments must require reasoning, disciplinary application, and traceable processes. Examples of allowed assistance reduce uncertainty and ensure consistency.

For providers, educational personalisation needs traceability, source support, accessibility, privacy, and output review. Optimising for interaction fluency may conflict with developing independent competence.

9. Conclusion

GenAI expands the possibilities of personalised learning through adaptable content, conversational tutoring, rapid feedback, accessibility support, and assistance for educators. The reviewed evidence also shows that these affordances are inseparable from risks of inaccurate output, bias, privacy loss, inequity, academic-integrity pressure, and reduced cognitive engagement. Current research is too heterogeneous and short-term to justify claims of automatic improvement. The most credible direction is a human-centred partnership in which GenAI provides scalable support while educators retain responsibility for pedagogical design, verification, relationships, and ethical judgement. With explicit AI literacy, assessment redesign, equitable access, and continuous evaluation, GenAI may strengthen personalised learning. Without these conditions, personalisation risks becoming efficient content generation without meaningful learning.

References

- [1] Rensfeldt, A. B., & Rahm, L. (2022). Automating teacher work? A history of the politics of automation and artificial intelligence in education. *Postdigital Science and Education*, 5(1), 25–43. <https://doi.org/10.1007/s42438-022-00344-x>
- [2] Mallikarjuna, B., & Chittamsetty, P. (2024). Generative artificial intelligence: Fundamentals and evolution. In *Generative AI: Current Trends and Applications* (pp. 3–17). https://doi.org/10.1007/978-981-97-8460-8_1
- [3] Renugadevi, R., Shobana, J., Arthi, K., Kalpana, A. V., Satishkumar, D., & Sivaraja, M. (2024). Real-time applications of artificial intelligence technology in daily operations. In *Advances in Computational Intelligence and Robotics Book Series* (pp. 243–257). <https://doi.org/10.4018/979-8-3693-2615-2.ch012>
- [4] Powell, W., & Kusuma-Powell, O. (2011). *How to teach now: Five keys to personalized learning in the global classroom*. ASCD.
- [5] Elmourabit, Z., Retbi, A., & El Faddouli, N. E. (2024). The impact of generative artificial intelligence on education: A comparative study. *European Conference on E-Learning*, 23(1), 470–476. <https://doi.org/10.34190/ecel.23.1.2975>
- [6] Granda, B. S., Inzhivotkina, Y., Ibáñez, F., & Gustavo, J. (2024). Educational innovation: Exploring the potential of generative artificial intelligence in cognitive schema building. *Educat Revista Electrónica de Tecnología Educativa*, 89, 44–63. <https://doi.org/10.21556/edutec.2024.89.3251>
- [7] Rawat, S., Mittal, S., Nehra, D., Sharma, C., & Kamboj, D. (2023). Exploring the potential of ChatGPT to improve experiential learning in education. *The 5th International Conference on Information Management & Machine Intelligence*, 83(1). <https://doi.org/10.1145/3647444.3647910>
- [8] Faccia, A., Ridon, M., Beebejaun, Z., & Mosteanu, N. (2023). Advancements and challenges of generative AI in higher educational content creation: A technical perspective. In *ICISE '23: Proceedings of the 2023 8th International Conference on Information Systems Engineering*. <https://doi.org/10.1145/3641032.3641055>
- [9] Mithi, J., Madzvamuse, S., Mbanje, S., & Lomahoza, S. (2024). Generative artificial intelligence and formative assessment: Perspectives from higher education in South Africa. *Proceedings of the 1st International Conference on Education Research*, 1(1), 449–458. <https://doi.org/10.34190/icer.1.1.3231>
- [10] Wu, D., Zhang, S., Ma, Z., Yue, X.-G., & Dong, R. K. (2024). Unlocking potential: Key factors shaping undergraduate self-

- directed learning in AI-enhanced educational environments. *Systems*, 12(9), 332. <https://doi.org/10.3390/systems12090332>
- [11] Gupta, N., Khatri, K., Malik, Y., Lakhani, A., Kanwal, A., Aggarwal, S., & Dahuja, A. (2024a). Exploring prospects, hurdles, and road ahead for generative artificial intelligence in orthopedic education and training. *BMC Medical Education*, 24(1). <https://doi.org/10.1186/s12909-024-06592-8>
- [12] Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behaviour*, 8(10), 1839–1850. <https://doi.org/10.1038/s41562-024-02004-5>
- [13] Chauhan, J. A., Bhatt, P. P., & Raval, D. K. (2024). The impact of generative AI in modern education system in India. In *Lecture Notes in Networks and Systems* (pp. 593–603). https://doi.org/10.1007/978-981-97-6684-0_48
- [14] Gabriel, S. (2024). Generative AI and educational (in)equity. *International Conference on AI Research*, 4(1), 133–142. <https://doi.org/10.34190/icaire.4.1.3153>
- [15] Yogi, M. K., Chowdary, Y. R., & Santhoshi, C. P. R. S. (2024). Impact of generative AI models on personalized learning and adaptive systems. In *Empowering Digital Education with ChatGPT* (pp. 83–97). <https://doi.org/10.1201/9781032716350-6>
- [16] Freise, L. R., & Ritz, E. (2024). Using ChatGPT to prepare for your dream job? In *Chapman and Hall/CRC EBooks* (pp. 142–155). <https://doi.org/10.1201/9781032716350-10>
- [17] Miao, J., Thongprayoon, C., Craici, I. M., & Cheungpasitporn, W. (2024). How to incorporate generative artificial intelligence in nephrology fellowship education. *Journal of Nephrology*, 37(9), 2491–2497. <https://doi.org/10.1007/s40620-024-02165-6>
- [18] Tang, V., Wong, L., Lam, H. Y., & Tang, Y. M. (2024). Amplifying learning and teaching effectiveness through generative artificial intelligence: A qualitative approach with case studies on supply chain and cold chain management. In *International Conference on Decision Science & Management* (pp. 264–269). <https://doi.org/10.1145/3686081.3686126>
- [19] Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- [20] Ilieva, G., Yankova, T., Klisarova-Belcheva, S., Dimitrov, A., Bratkov, M., & Angelov, D. (2023). Effects of generative chatbots in higher education. *Information*, 14(9), 492. <https://doi.org/10.3390/info14090492>
- [21] Arslan, B., Lehman, B., Tenison, C., Sparks, J. R., López, A. A., Gu, L., & Zapata-Rivera, D. (2024). Opportunities and challenges of using generative AI to personalize educational assessment. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/frai.2024.1460651>
- [22] Gupta, S., Dharamshi, R. R., & Kakde, V. (2024b). An impactful and revolutionized educational ecosystem using generative AI to assist and assess the teaching and learning benefits, fostering the post-pandemic requirements. In *2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*. <https://doi.org/10.1109/ic-etite58242.2024.10493370>
- [23] Li, Y., Ji, W., Liu, J., & Li, W. (2024). Application of generative artificial intelligence technology in customized learning path design: A new strategy for higher education. In *2024 International Conference on Interactive Intelligent Systems and Techniques (IIIST)* (pp. 567–573). <https://doi.org/10.1109/iist62526.2024.00099>
- [24] Gupta, A. (2024). Impact of generative AI in transforming higher education pedagogy. In *Advances in Business Information Systems and Analytics Book Series* (pp. 285–300). <https://doi.org/10.4018/979-8-3693-0815-8.ch017>
- [25] Zhang, Y., & Miao, M. (2024). The roles and functions of foreign language teachers of China in the context of generative artificial intelligence. In *2024 International Conference on Intelligent Education and Computer Technology* (pp. 584–586). <https://doi.org/10.1145/3687311.3687415>
- [26] Roy, A. D., Das, D., Dasgupta, S., & Roy, R. D. (2024). The use of generative artificial intelligence (AI) in teaching and assessment of postgraduate students in pathology and microbiology. *Indian Journal of Microbiology Research*, 11(3), 140–146. <https://doi.org/10.18231/ij.jmr.2024.027>
- [27] Pereira, C. S., Ferreira, M. J., Sobral, S. R., Durão, N., & Moreira, F. (2024). Enhancing higher education in Portugal: Leveraging generative artificial intelligence for learning-teaching process. *European Conference on E-Learning*, 23(1), 289–298. <https://doi.org/10.34190/ecel.23.1.2503>
- [28] Silva, P. (2015). Davis' technology acceptance model (TAM) (1989). In *Information Seeking Behavior and Technology Adoption* (pp. 205–219). <https://doi.org/10.4018/978-1-4666-8156-9.ch013>
- [29] Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, 5(6), 66–70.
- [30] Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, 11(2), 87–95. <https://doi.org/10.1080/00461527409529129>
- [31] Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. <https://doi.org/10.2307/j.ctvjf9vz4>
- [32] Hannes, K., & Claes, L. (2007). Learn to read and write systematic reviews: The Belgian Campbell Group. *Research on Social Work Practice*, 17(6), 748–753. <https://doi.org/10.1177/1049731507303106>
- [33] Johnson, B. R., De Li, S., Larson, D. B., & McCullough, M. (2000). A systematic review of the religiosity and delinquency literature. *Journal of Contemporary Criminal Justice*, 16(1), 32–52. <https://doi.org/10.1177/1043986200016001003>
- [34] Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences*. Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470754887>
- [35] Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7). <https://doi.org/10.1371/journal.pmed.1000097>
- [36] Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 1–10. <https://doi.org/10.1186/1471-2288-8-45>
- [37] Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- [38] Pesovski, I., Santos, R., Henriques, R., & Trajkovic, V. (2024). Generative AI for customizable learning experiences. *Sustainability*, 16(7), 3034. <https://doi.org/10.3390/su16073034>
- [39] Rathika, P., Yamunadevi, S., Ponni, P., Parthipan, V., & Anju, P. (2024). Developing an AI-powered interactive virtual tutor for enhanced learning experiences. *International Journal of Computational and Experimental Science and Engineering*, 10(4). <https://doi.org/10.22399/ijcesen.782>