

Integrating Artificial Intelligence with Gamification in Medical Education: A Pedagogically Grounded Framework and Critical Review

Ambadasu Bharatha¹, Nkemcho Ojeh¹, Michael H Campbell²,
Kandamaran Krishnamurthy^{2,3}, Vijay Prasad Sangishetti⁴, Lalitha Bhuvanagiri⁵, Subir Gupta¹,
Sayeeda Rahman⁶, Sudharshan Reddy⁷, Md Anwarul Azim Majumder⁸

¹Department of Preclinical Sciences, Faculty of Medical Sciences, University of the West Indies, Bridgetown, Barbados; ²Department of Clinical Sciences, Faculty of Medical Sciences, University of the West Indies, Bridgetown, Barbados; ³Pediatric Intensive Care Unit, Queen Elizabeth Hospital, Bridgetown, Barbados; ⁴Department of Pharmacology, SRVS Govt Medical College, Shivpuri, MP, India; ⁵TRR Institute of Medical Sciences, Patancheru, Hyderabad, India; ⁶Department of Pharmacology, Bridgetown International University School of Medicine, Bridgetown, Barbados; ⁷Department of Pharmacology, Shri BM Patil Medical College Hospital and Research Centre BLDE (DU), Vijayapura, KA, India; ⁸Department of Clinical Sciences, Bridgetown International University School of Medicine, Bridgetown, Barbados

Correspondence: Md Anwarul Azim Majumder, Professor of Medical Education & Dean of Clinical Sciences, Bridgetown International University School of Medicine, Bridgetown, Barbados, Email clinicaldean@biuuniversity.org; Michael H Campbell, Department of Clinical Sciences, Faculty of Medical Sciences, University of the West Indies, Bridgetown, Barbados, Email michael.campbell@uwi.edu

Abstract: Digital learning technologies have transformed medical education, with artificial intelligence (AI) and gamification emerging as two of the most active areas of innovation. While each has demonstrated value independently, their integration offers distinctive potential to personalise learning, sustain engagement, and produce durable educational outcomes. Yet the convergence remains empirically disjointed and theoretically underdeveloped. This pedagogically grounded critical review synthesises the evidence at the crossroads of AI and gamification in medical and health professional education, drawing on randomised trials, scoping reviews, meta-analyses, and case studies from PubMed, DOAJ, ERIC, and Web of Science. We propose an operational definition of AI-enhanced gamification and introduce an integration matrix linking five AI methods (reinforcement learning, Bayesian learner modelling, natural language processing, computer vision, recommender systems) to specific gamification elements and to learning mechanisms grounded in Self-Determination Theory, Flow Theory, constructivism, Vygotsky's Zone of Proximal Development, connectivism, the TPACK model, and the Behaviour Change Technique taxonomy. We map applications across health literacy, mental health psychoeducation, rehabilitation, and medical education, supported by ten real-world examples. We identify challenges in theoretical grounding, outcome measurement, validation, algorithmic bias, reproducibility, equity, and regulation, and close with a prioritised, feasibility-tagged research agenda for advancing AI-enhanced gamification as a credible digital learning innovation in medical education.

Keywords: artificial intelligence, gamification, digital learning innovation, educational technology, adaptive learning, instructional design, personalised learning, TPACK, medical education

Introduction

Digital learning technologies have fundamentally transformed teaching and learning across classrooms, clinical training environments, and the workplace.^{1,2} Within this broader shift, medical and health professional education has become a particularly active focus of digital innovation. The rising global burden of chronic disease has created unprecedented demand not only for clinical services but also for effective educational interventions that empower clinicians, patients, families, and communities to manage illness and make informed health decisions.^{3,4} At the same time, medical education itself must produce graduates capable of adapting to rapidly changing clinical environments, embracing emerging technologies, and engaging in lifelong learning.⁵

Two technologies are positioned to meet these challenges. The first is artificial intelligence (AI), a broad field encompassing machine learning, deep learning, natural language processing, and reinforcement learning, which offers transformative potential for personalising learning, tailoring instructional content, and providing real-time formative feedback.^{6,7} The second is gamification, defined by Deterding et al as the use of game-design elements in non-game contexts, widely used in educational settings to enhance motivation, engagement, and behaviour change.^{8,9} Educationally, gamification aligns with constructivist principles, positioning the learner as an active builder of knowledge through interaction, feedback, and progressive challenge.¹⁰

Each technology has built a substantial individual evidence base. AI-driven intelligent tutoring systems have improved learning across domains including clinical reasoning,¹¹ and hundreds of gamified health applications have demonstrated positive effects on engagement and knowledge retention.^{12,13} Their intentional combination offers further pedagogical potential: AI systems that dynamically tailor gamified learning experiences to the individual learner — adjusting difficulty, reward schedules, feedback timing, and narrative — using continuous models of knowledge state and motivation.

Despite growing interest, the convergence of AI and gamification in medical education remains empirically disjointed. A 2024 scoping review identified just 16 studies exploring AI-integrated game-based applications in healthcare,¹⁴ most being small-scale pilots with limited grounding in educational theory or learning outcomes.¹⁵ This gap is striking given the well-developed theoretical foundations available — the TPACK model,¹⁶ the Cognitive Theory of Multimedia Learning,¹⁷ and Merrill's First Principles of Instruction¹⁸ — which offer robust guidance for designing and assessing technology-enhanced learning.

Drawing on randomised trials, scoping reviews, meta-analyses, and case studies identified through PubMed, DOAJ, ERIC, and Web of Science, this pedagogically grounded critical review synthesises the evidence at the crossroads of AI and gamification in medical and health professional education. Moving beyond technological description, it interrogates the pedagogical foundations, instructional design, and measurable learning outcomes of AI-enhanced gamified interventions, and is organised around the structure summarised in [Figure 1](#): from the central problem of disjointed integration, through the theoretical foundations, AI methods, and integration framework that anchor our analysis, to the real-world evidence, seven integration challenges, and seven-priority research agenda that follow.

Educational and Psychological Foundations of Gamification

Defining Gamification Within Educational Discourse

Gamification entered mainstream learning research in 2010, having originated in earlier studies of serious games, game-based learning, and persuasive technology.⁸ Gamification is frequently defined as the use of game-design elements in non-games, although more precise definitions exist, especially in educational research where scholars have differentiated between structural gamification (using game elements with non-game content) and content gamification (integrating content into game-like narratives and environments).⁹ Werbach and Hunter also described game dynamics, mechanics, and components, offering a vocabulary of design that has since gained prominence in educational technology research.¹⁹ In health-related education, gamification can be distinguished from serious games (full games developed for an educational purpose but not primarily for entertainment) and exergames (games that involve physical activity and interaction), because each represents a particular pedagogical strategy with different implications for the design and evaluation of learning.⁵

Sardi et al provided an early systematic review²⁰ offering a comprehensive mapping of gamified e-health applications. Based on analysis of 46 studies, the authors found that most gamified health interventions used points, levels, and feedback loops; however, very few described a clear educational or theoretical rationale for the key design elements, and even fewer evaluated which specific components contributed to observed learning. This theory–practice divide is a consistent and significant weakness that reflects the broader tendency of educational technology research to adopt technologies without sufficient theoretical justification.²¹

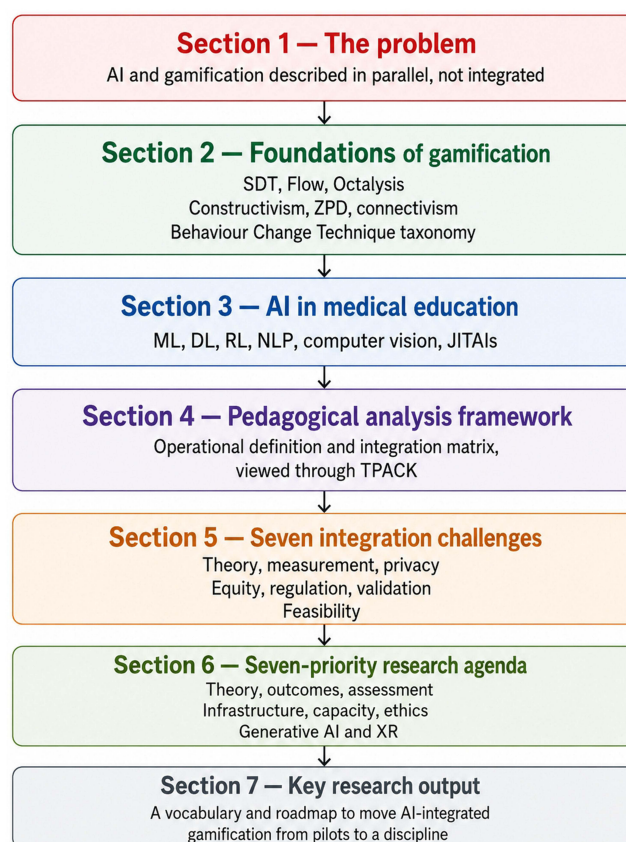


Figure 1 Graphical summary of the review. [Credit: This image was generated using ChatGPT Images 2.0, OpenAI, 27 June 2026, <https://chatgpt.com/>].

Abbreviations: AI, artificial intelligence; SDT, Self-Determination Theory; ZPD, Zone of Proximal Development; ML, machine learning; DL, deep learning; RL, reinforcement learning; NLP, natural language processing; JITAs, just-in-time adaptive interventions; TPACK, Technological Pedagogical Content Knowledge; XR, extended reality.

Theoretical Foundations of Gamification: Implications for Medical Education

Several psychological and educational theories explain how gamification improves teaching, learning, and assessment. Self-Determination Theory (SDT) holds that intrinsic motivation depends on autonomy, competence, and relatedness.²² Gamification operationalises all three: optional quests and pacing preserve autonomy, calibrated difficulty sustains competence, and team challenges or leaderboards cultivate relatedness, with autonomy-supportive designs consistently linked to deeper learning and durable behaviour change in health contexts.²³ Csikszentmihalyi's Flow Theory predicts deep engagement when challenge matches skill,²⁴ gamified systems pursue this balance through tiered difficulty and immediate feedback, producing the prolonged time on task and improved retention associated with flow states.²⁵ Chou's Octalysis framework translates motivation into eight design drives,²⁶ offering a planning vocabulary for educators, although empirical validation in clinical learner groups remains limited.²⁷

Beyond motivational theory, constructivism treats learners as active builders of knowledge through experience and social interaction,¹⁰ which gamified problem scenarios and simulations operationalise by replacing passive content delivery with decision-making and reflection. Vygotsky's Zone of Proximal Development further informs assessment design: scaffolded tasks, withdrawn as competence grows, allow educators to measure both supported and independent performance.²⁸ Connectivism locates knowledge in distributed networks,²⁹ supporting gamified platforms that build peer communities and shared learning artefacts, particularly in chronic disease self-management. Finally, the Behaviour Change Technique (BCT) taxonomy provides a standardised vocabulary linking gamification mechanics to evidence-based change techniques, enabling pedagogically transparent design, planned implementation, and fidelity-coded assessment.³⁰ Table 1 summarises these theoretical foundations together with their implications for teaching, learning, and assessment in AI-enhanced gamification.

Table 1 Psychological and Educational Theories Underpinning Gamification in Digital Learning

Theory	Short Description	Implications for Teaching, Learning, and Assessment
Self-Determination Theory ^{22,23}	Intrinsic motivation needs autonomy, competence, and relatedness.	Offer choice and pacing; tune difficulty to skill; add social play; assess with SDT subscales.
Flow Theory ^{24,25}	Deep engagement occurs when challenge matches skill.	Use AI-driven dynamic difficulty; target 70–80% success; track engagement traces.
Octalysis Framework ^{26,27}	Eight motivational drives guide game-mechanic design.	Use the eight drives as a design checklist; personalise drives by learner profile; validate against learning outcomes.
Constructivism ¹⁰	Learners build knowledge through experience, reflection, and interaction.	Use problem-based scenarios and simulations; assess with performance and case tasks, not recall alone.
Zone of Proximal Development ²⁸	Learning happens just beyond independent ability, with scaffolded support.	Use adaptive hints and examples; withdraw support as competence grows; assess with and without scaffolds.
Connectivism ²⁹	Knowledge lives in networks; learning is making connections.	Build peer communities and shared knowledge artefacts; assess source evaluation and network participation.
Behaviour Change Technique taxonomy ³⁰	A standard vocabulary of 93 behaviour-change techniques.	Map each game element to a specific BCT; report BCTs used; evaluate intervention fidelity by BCT coding.

Note: References as numbered in the main manuscript.

Abbreviations: BCT, behaviour change technique; SDT, Self-Determination Theory.

Tensions Among Competing Theoretical Frameworks

Although these theories are often invoked together, they are not always internally consistent. The most common gamification elements — points, badges, leaderboards — are operant-conditioning artefacts rooted in behaviourism; yet they are routinely justified by appeal to Self-Determination Theory, which predicts that salient extrinsic rewards undermine intrinsic motivation (the overjustification effect). Reinforcement-learning-driven adaptive systems, by design, choose what content the learner sees next, which sits uneasily with SDT’s autonomy axiom. And the gamification literature routinely conflates engagement metrics (time on task, completion rates) with learning outcomes, even though multimedia-learning research has shown the two dissociate. Designers committed to constructivist or self-determination principles should treat extrinsic reward elements, opaque adaptation, and engagement-only assessment with caution rather than as defaults. We return to these tensions in Section 5 when we discuss the methodological and ethical challenges of AI–gamification integration.

Adoption of Gamification in Medical and Health Professional Education

In undergraduate and postgraduate medical education, gamification has played a prominent role in teaching clinical reasoning, procedural skills, patient communication, pharmacology, and interprofessional collaboration.⁵ Wang et al (2024) conducted a Delphi study in which expert panellists highlighted the importance of several gamification features for medical training: clear training objectives and challenges, personalisation options, feedback, freedom of choice, and the possibility of learning through failure.³¹ These elements reflect well-established principles of instructional design, indicating that gamification can be grounded in evidence-based teaching practice.

The future of medical education is likely to include the increasing adoption of AI-adaptive learning platforms. Systems that monitor learner progress through a sequence of gamified learning modules and identify knowledge gaps through Bayesian inference or item response theory, while sequencing challenges based on the principles of spaced repetition and interleaving, can enhance learning effectiveness and long-term retention.³² Gamified training environments based on AI-driven feedback loops have demonstrated the ability to decrease the learning curve for procedural competency in surgical simulation. Despite these promising findings, rigorous trials comparing AI-enhanced modalities with traditional training methods remain limited. Combining AI with simulation-based education, virtual patients, and team-based learning is promising for developing immersive and adaptive learning experiences that build not only knowledge but also clinical judgement, communication, and professional identity. Table 2 summarises ten recent real-world examples that document the impact of gamification across medical and health professional education, spanning

Table 2 Ten Recent Real-World Examples of Gamification with Documented Impact on Medical and Health Professional Education

#	Intervention	Learners	Design	Reported Impact
1	Septris — sepsis serious game (Stanford) ³³	156 medical learners	Pre/post study	Significant gains in sepsis recognition and management ($P < 0.001$); 85% would recommend
2	Game-based e-learning — phase-contrast urinalysis (Freiburg) ³⁴	145 third-year medical students	RCT vs script-based instruction	Game-based group: superior cognitive outcomes and higher learner enjoyment
3	Kahoot! — formative assessment ³⁵	Medical students (Malaysia)	Phenomenological study	Participation rose from <50% to >90%; perceived as motivating and engaging
4	Kahoot! — medication knowledge in CME ³⁶	128 nurses	Quasi-experimental	Significant gains in medication knowledge, performance, and satisfaction vs lecture ($P < 0.001$)
5	VR laparoscopic simulators ³⁷	12 RCTs / 241 participants	Systematic review	Operation time reduced 17–50%; grade IA–IIB evidence
6	VR cholecystectomy training ³⁸	Surgical trainees	Meta-analysis of RCTs	Improved OSATS performance; ~8.4 min reduction in task time
7	Educational escape rooms ³⁹	Medical students	Systematic review and meta-analysis	Knowledge SMD 0.84; teamwork SMD 4.91; learning satisfaction MD 0.36; reduced anxiety
8	Immersive VR/AR anatomy ⁴⁰	24 included studies	Systematic review and meta-analysis	Anatomy knowledge SMD 0.58 (95% CI 0.22–0.95); improved spatial understanding
9	Gamified flipped classroom ⁴¹	128 nursing students	RCT (6 weeks)	Higher skills competency, knowledge, and self-confidence (all $P < 0.05$)
10	Digital serious games in nursing ⁴²	11 included studies / 8 in meta-analysis	Systematic review and meta-analysis of RCTs	Significant gains in knowledge (Hedges' $g = 0.75$), confidence (Hedges' $g = 0.73$), and performance (Hedges' $g = 0.49$); low risk of bias

Abbreviations: AR, augmented reality; CI, confidence interval; CME, continuing medical education; MD, mean difference; OSATS, Objective Structured Assessment of Technical Skill; RCT, randomised controlled trial; SMD, standardised mean difference; VR, virtual reality.

serious games for clinical reasoning, virtual-reality simulators for procedural training, formative-assessment platforms, and team-based escape rooms.

Artificial Intelligence in Medical Education: A Revolution in the Making

Artificial intelligence (AI) refers to computational systems that perform tasks normally requiring human cognition — learning from data, recognising patterns, reasoning under uncertainty, and generating language. Its principal subfields include machine learning (ML), in which algorithms improve from data without explicit programming;⁶ deep learning (DL), which uses multi-layered neural networks to model complex, high-dimensional inputs such as images, movement, and speech;⁴³ reinforcement learning (RL), in which agents learn optimal sequential actions through reward feedback;⁴⁴ and natural language processing (NLP), including large language models (LLMs), which interpret and generate human language.⁴⁵

These methods are reshaping medical education across four dimensions. First, personalisation: supervised ML models predict learners at risk of disengagement, profile knowledge gaps, and adapt content difficulty and pacing in real time.^{43,46} Second, real-time formative assessment: convolutional neural networks evaluate procedural and motor skills — surgical technique, physiotherapy form, communication behaviours — delivering immediate corrective feedback that previously required constant expert supervision.⁴⁷ Third, adaptive sequencing: RL agents optimise the order and timing of learning material, as demonstrated in the DIAMANTE trial, where RL-driven message tailoring produced a 19% step-count increase relative to control.⁴⁸ Just-in-time adaptive interventions (JITAI) extend this principle, delivering instruction at the precise moment of learner need through sensors and context-aware models.^{49–51} Fourth, conversational pedagogy: NLP-driven tutors and LLM-based virtual patients apply Socratic questioning, scaffolded explanation, and clinical-reasoning simulation, while generative AI dynamically authors narrative scenarios and literacy-matched patient education content.⁵²

Together, these capabilities transform medical training from fixed curricula into responsive, data-driven learning environments. Realising this promise requires rigorous attention to algorithmic transparency, validation in diverse learner populations, and content-fidelity safeguards — challenges addressed in later sections.

A Pedagogical Analysis Framework: Connecting AI-Enhanced Gamification to Educational Models

Terms such as *AI-enhanced gamification*, *intelligent gamification*, and *adaptive serious games* are used inconsistently across the literature. For this review, we propose the following operational definition:

AI-enhanced gamification is a class of digital learning intervention in which game-design elements (eg, challenge calibration, quest sequencing, narrative content, feedback timing, reward schedules, social matching) are dynamically configured by machine-learning systems modelling the individual learner's knowledge state, motivation, and context, to produce measurable, theory-grounded learning outcomes.

This review argues that the AI–gamification literature must move beyond engagement metrics and clinical endpoints to engage seriously with established educational frameworks that ensure quality of learning, valid assessment, and transferability of knowledge and skills.^{16,21} To make the AI–gamification crosstalk concrete rather than asserted, we propose an integration matrix (Figure 2) that operationalises the definition above by linking five common AI methods to specific gamification elements they modify and to the learning mechanisms those modifications recruit. Reinforcement learning enables sequential decision-making,⁴⁴ manifesting as dynamically calibrated difficulty curves that recruit Csikszentmihalyi's flow states by maintaining an optimal challenge–skill ratio.^{24,25} Bayesian learner models estimate the knowledge state,⁵³ allowing quest sequencing that respects Vygotsky's Zone of Proximal Development.²⁸ Natural language processing and large language models enable narrative branching personalised to the learner's profile,^{45,54} supporting the autonomy axis of Self-Determination Theory.^{22,23} Convolutional neural networks and pose-estimation models assess motor performance,^{47,55} delivering just-in-time augmented feedback aligned with motor-learning theory.⁵⁴ Recommender systems perform social matching,⁴³ recruiting the relatedness axis of SDT²² and connectivist principles of networked knowledge.²⁹ Each row is a falsifiable claim: an AI component that merely sits alongside a game element, without modifying it, does not constitute AI-enhanced gamification.

This synergy is best understood through Mishra and Koehler's TPACK framework,¹⁶ which holds that effective technology-enhanced teaching requires integrated technological, pedagogical, and content knowledge. Most current AI-gamified interventions display strong technological knowledge but inadequate pedagogical and content grounding: gamification elements are chosen for engagement rather than learning,^{14,20} AI adaptation optimises usage rather than mastery,⁴³ and assessment captures compliance rather than knowledge construction.⁵⁶ Closing this TPACK gap requires interdisciplinary teams uniting learning scientists, instructional designers, AI engineers, and game developers from the design stage onward.

In summary, Figure 2 visualises the integration matrix and offers a falsifiability criterion: any candidate AI-gamified intervention should map cleanly onto at least one row, naming the AI method, the gamification element it modifies, and the learning mechanism it recruits.

Integration of AI with Gamification in Medical Education: Opportunities, Barriers, and Ethical Considerations

The integration of artificial intelligence with gamification offers a powerful pathway to transform medical education, yet realising this potential requires confronting a set of pedagogical, methodological, ethical, and regulatory challenges that the current evidence base has only begun to address. This section synthesises these issues and the safeguards needed for responsible implementation; Figure 3 summarises the seven integration challenges and their corresponding safeguards discussed below.

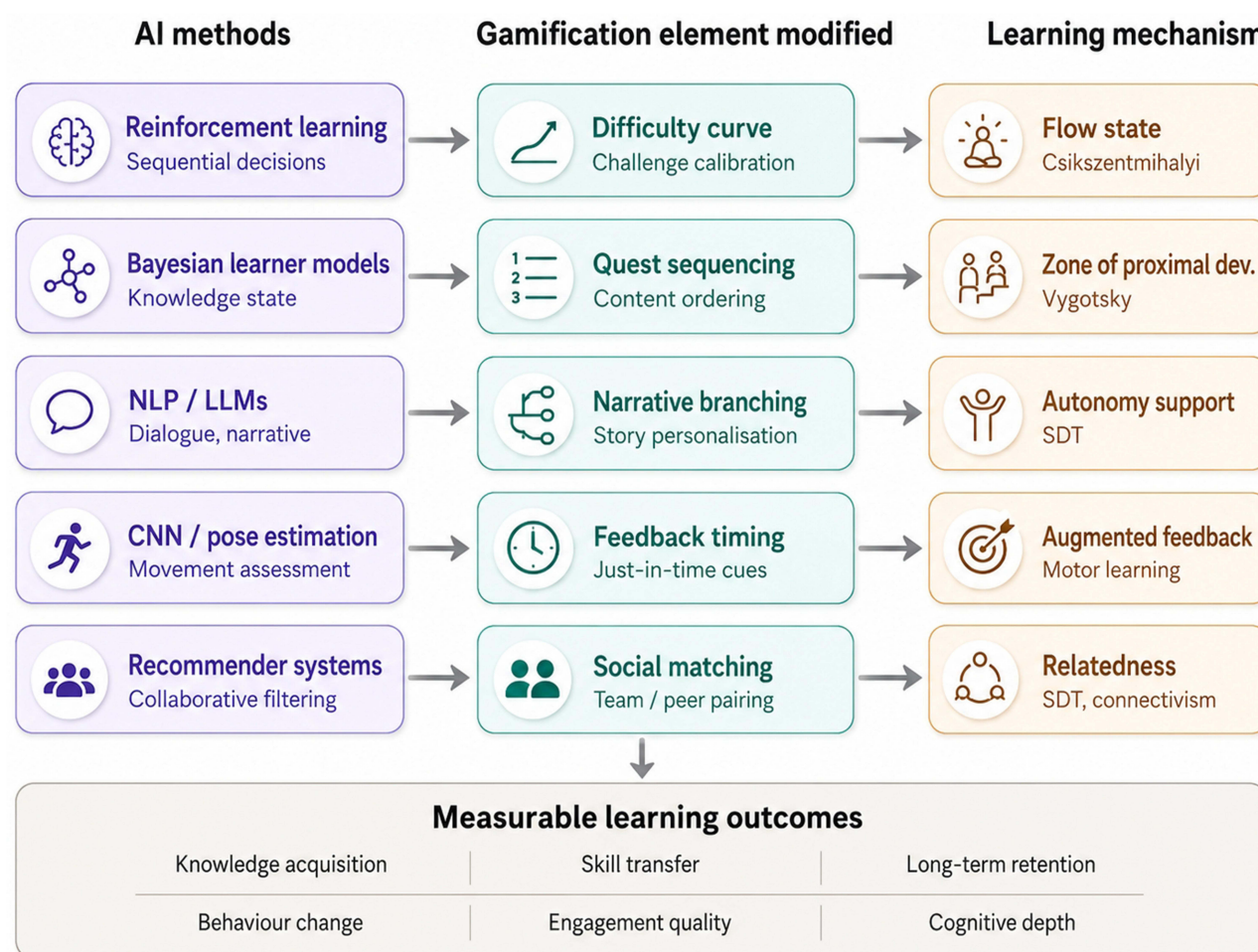


Figure 2 AI-gamification integration matrix. [Credit: This image was generated using ChatGPT Images 2.0, OpenAI, 27 June 2026, <https://chatgpt.com/>]. AI methods (left) modify specific gamification elements (centre), which in turn recruit learning mechanisms grounded in established educational theory (right), together producing measurable learning outcomes in medical education.

Bridging the Theory–Practice Gap in Instructional Design

The most prominent obstacle to integrating AI with gamification in medical education is the persistent disconnect between educational theory and intervention design. Despite the rich theoretical foundations of educational technology, the learning sciences, and instructional design, most studies in the current evidence base provide little pedagogical justification for their design decisions. Tolks et al found that 88% of studies were non-randomised pilots with minimal theoretical basis,¹⁴ and Nishi et al reported wide variability in outcomes and intervention designs, complicating evidence synthesis.⁵⁷ Without explicit alignment between gamification elements, AI components, and learning goals, the field cannot link interventions to educational outcomes. Closing this gap requires applying established instructional design models - the ADDIE model (Analysis, Design, Development, Implementation, Evaluation) and Merrill's First Principles of Instruction, which emphasise problem-centred learning, prior-knowledge activation, demonstration, application, and integration¹⁸ - from the earliest stages of AI-gamified curriculum design.

Methodological Rigour and Educational Outcome Measurement

A second barrier to integration concerns how learning is measured. Most studies evaluate engagement (time-on-task, completion rates) or behavioural surrogates (step counts, medication adherence) rather than direct educational outcomes such as knowledge acquisition, skill development, transfer, and long-term retention. Disentangling the independent contributions of AI personalisation and gamification mechanics is further complicated when adaptive difficulty,

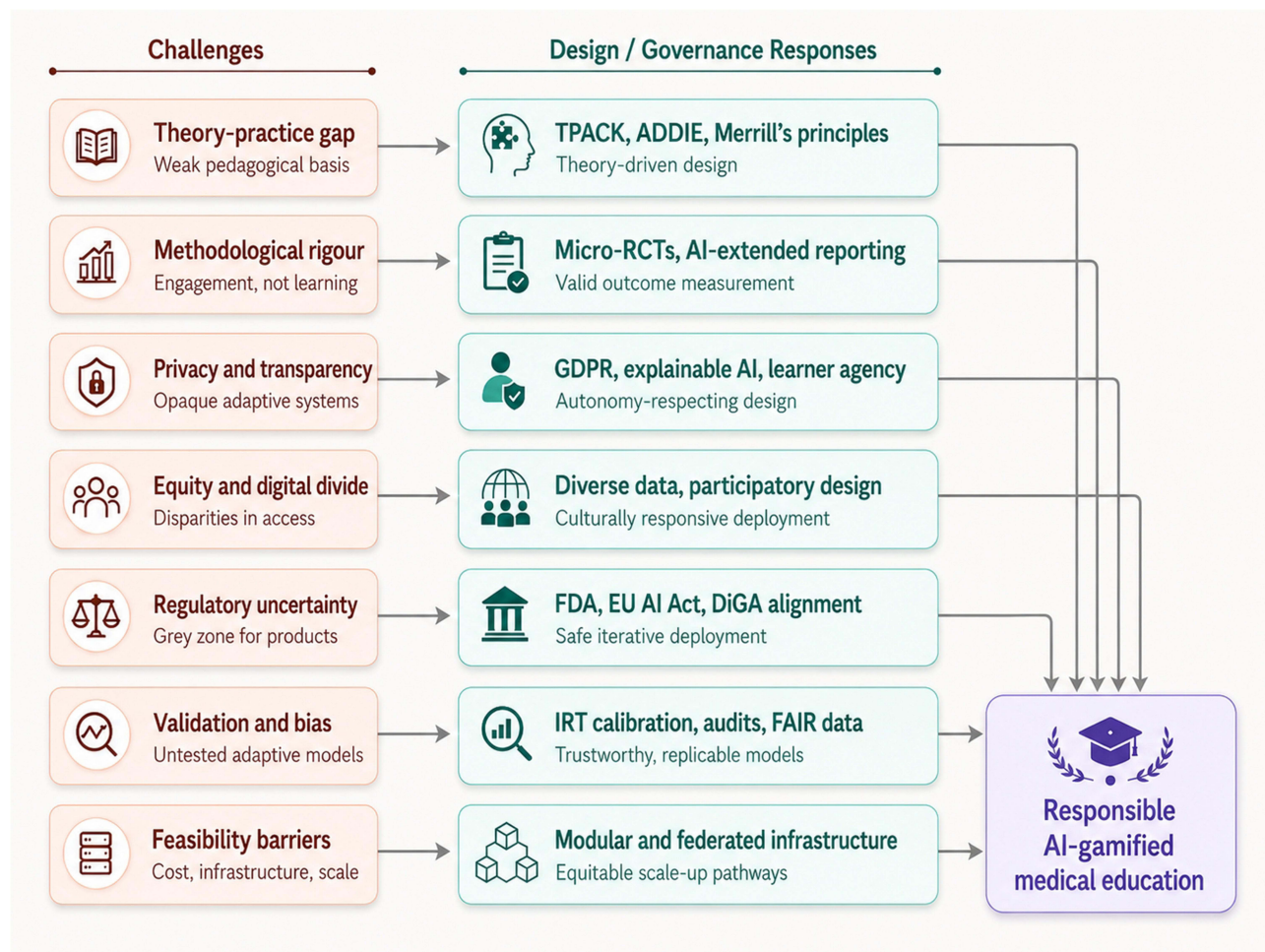


Figure 3 Integration challenges and safeguards. [Credit: This image was generated using ChatGPT Images 2.0, OpenAI, 27 June 2026, <https://chatgpt.com/>].

personalised rewards, and social matching are deployed together. Micro-randomised trials, which randomly allocate intervention components at single decision points, offer an appealing methodology but remain uncommon in this field.⁵⁸ Reporting standards also lag behind the technology: the CONSORT-EHEALTH checklist provides useful guidance for digital health trials but does not address adaptive AI-based educational interventions,⁵⁹ and an AI-specific extension is overdue.

Data Privacy, Algorithmic Transparency, and Learner Autonomy

AI-gamified medical education is intrinsically data-intensive, relying on continuous monitoring of learner behaviour, performance, and context. This raises substantial privacy concerns, particularly when commercial vendors handle sensitive learner and health data. The European GDPR and analogous frameworks offer guidance, but meaningful informed consent is difficult when adaptive systems themselves change over time.⁶⁰ Algorithmic transparency further complicates learner autonomy: reinforcement-learning agents may present different content to learners with similar profiles depending on exploration–exploitation processes, yet neither learner nor educator can readily explain why a particular item was delivered at a particular moment.⁴⁴ The emerging field of explainable AI (XAI) promises to clarify these decisions, but applications to gamified medical learning remain limited.

Educational Equity and the Digital Divide

The integration of AI with gamification carries a real risk of exacerbating existing disparities. Access to suitable devices, reliable internet, and the digital literacy required to benefit from adaptive platforms is unevenly distributed; older learners,

those in low-resource settings, and learners with limited prior technology exposure are most likely to be left behind.⁶¹ The DIAMANTE trial's deliberate inclusion of multilingual learners from public primary care offers a valuable model of equitable design but remains the exception.⁴⁸ Cultural responsiveness matters as well: norms around play, competition, and collaboration vary across contexts, and AI systems trained predominantly on data from young, technology-literate, Western learners may generate strategies that translate poorly elsewhere.⁶² Equitable integration requires diverse training datasets, participatory design with end users, and validation across the social and cultural contexts of deployment.

Regulatory Considerations for AI-Gamified Educational Products

The regulatory landscape is evolving rapidly. The US FDA's Predetermined Change Control Plan guidance aims to permit iterative improvement of AI-enabled software while preserving safety and effectiveness assurances,⁶³ but many gamified educational products occupy a grey zone between wellness apps and medical devices.⁶⁴ The EU AI Act and Medical Devices Regulation classify AI-driven medical devices as high-risk, requiring substantial documentation and human oversight, and Germany's Digital Healthcare Act (DiGA) has established a reimbursement pathway for evidence-based digital therapies prescribed by physicians.⁶⁵ Key questions remain: how should continuously learning algorithms be re-evaluated post-approval, and what evidence thresholds should apply to gamified educational platforms positioned as adjuncts to formal training? As regulatory frameworks mature, alignment between educational, clinical, and software-as-a-medical-device standards will be essential for safe, scalable integration.

Model Validation, Algorithmic Bias, and Reproducibility in Medical-Education Datasets

Beyond structural and ethical concerns, AI-gamified medical education raises three technical challenges.

- Validation: adaptive systems change with each learner, so conventional psychometric metrics do not apply; item response theory (IRT) calibration of stealth assessments against MCQ banks, OSCEs, and entrustable professional activity ratings, together with prospective external validation and micro-randomised trials, are needed.^{56,58}
- Bias: training data predominantly drawn from well-resourced, English-language institutions can produce models that under-serve learners and patients in other contexts, requiring explicit fairness criteria and ongoing audits.^{61,62}
- Reproducibility: vague algorithmic descriptions and inaccessible datasets prevent independent replication; adoption of FAIR principles and an AI-specific extension of CONSORT-EHEALTH would substantially improve the field.^{14,59}

Feasibility: Cost, Infrastructure, and Scalability Barriers

Translating AI-gamified medical education from pilot to mainstream curriculum is constrained by tangible feasibility barriers. *Cost* spans software licensing, content authoring, GPU compute for adaptive models, and faculty time for design and maintenance - recurring expenses that small or resource-limited institutions cannot easily absorb. *Infrastructure* requirements include reliable high-bandwidth internet, learner devices capable of running immersive content, learning-management-system integration, and secure data pipelines for adaptive personalisation; these are unevenly available across medical schools globally.⁶¹ *Scalability* is further constrained by the need for institution-specific content adaptation, regulatory variation across jurisdictions,^{63,64} and faculty TPACK capacity.¹⁶ Modular, reusable AI-gamification components, federated infrastructure shared across institutions,⁶⁶ and tiered “good–better–best” implementation models offer practical pathways to equitable scale-up.

Together, these seven challenges define the landscape of integration risk in AI-enhanced gamification for medical education. [Figure 3](#) visualises each challenge alongside the safeguard required to address it, converging on the responsible deployment of AI-gamified interventions.

Future Directions and Research Agenda: AI Integration with Gamification in Medical Education

The integration of artificial intelligence with gamification offers a credible pathway to more personalised, engaging, and pedagogically rigorous medical education, but realising this potential requires deliberate, prioritised research. The recommendations below identify seven priority directions for the field, visualised in [Figure 4](#).

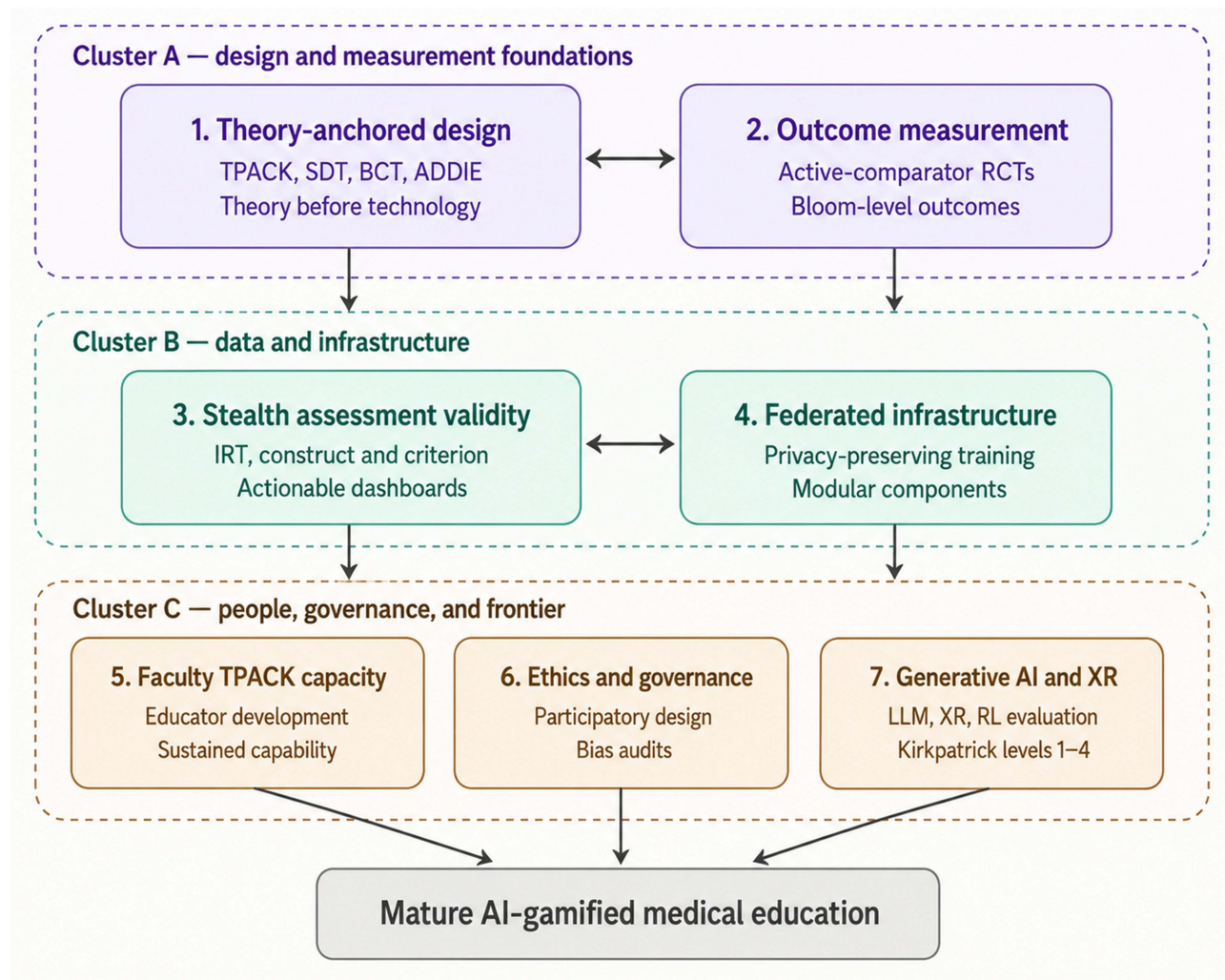


Figure 4 Prioritised research agenda. [Credit: This image was generated using ChatGPT Images 2.0, OpenAI, 27 June 2026, <https://chatgpt.com/>].

Anchor AI-Gamified Medical Education in Established Instructional Design

The most consistent weakness of the current evidence base is the absence of explicit pedagogical grounding. Future AI-gamified interventions for medical and health professional education should be designed from established educational frameworks — Self-Determination Theory, constructivism, TPACK, and the Behaviour Change Technique taxonomy — rather than retrofitted to them after deployment.^{16,18,30} Systematic instructional design models such as ADDIE (Analysis, Design, Development, Implementation, Evaluation) and the Successive Approximation Model (SAM) provide structured approaches for translating theory into AI-gamified curricula. Development teams should routinely include learning scientists and instructional designers alongside AI engineers and game developers, with pedagogical objectives specified before technological decisions are made.

Strengthen Educational Outcome Measurement in AI-Gamified Trials

The field urgently needs adequately powered randomised controlled trials that move beyond engagement metrics to validated educational outcomes — knowledge acquisition, skills development, transfer, long-term retention, and clinical performance.^{14,67} Trials should use active rather than passive comparators (non-adaptive gamification versus AI-only adaptive instruction) to disentangle the independent contributions of AI personalisation and gamification mechanics. Bloom's revised taxonomy provides a framework for assessing depth of cognitive learning, and reporting should follow CONSORT-EHEALTH⁵⁹ with an AI-specific extension addressing adaptive algorithm specification, reward functions, and learner-model architecture.

Develop and Validate Learning Analytics and Stealth Assessment

AI-gamified medical education generates rich behavioural and performance data, but the validity of these traces as indicators of genuine learning remains under-examined.^{56,68} Future research should validate stealth-assessment tools against external knowledge instruments using item response theory; demonstrate construct validity (measuring learning rather than game mastery), criterion validity (predicting real-world clinical performance), and measurement invariance across diverse learner groups; and develop learning-analytics dashboards that give educators and trainees actionable, interpretable feedback rather than opaque engagement scores.

Advance Privacy-Preserving and Federated Learning Architectures

AI-gamified medical education depends on continuous learner data, raising legitimate privacy concerns particularly for trainees whose performance data could affect career trajectories. Federated learning and other privacy-preserving machine-learning approaches allow personalisation algorithms to be trained across institutional boundaries without centralising sensitive data.⁶⁶ Future research should evaluate the feasibility of multi-institutional federated training for medical education AI, develop reusable modular AI-gamification components that lower development costs across schools, and establish data-governance frameworks specific to educational rather than clinical contexts.

Invest in Faculty Development and TPACK Capacity-Building

Technology will not transform medical education without educators capable of designing, implementing, and evaluating it. TPACK-capable faculty — those who can integrate technological knowledge with pedagogical and content expertise — are currently rare in medical schools.¹⁶ Research should evaluate structured faculty-development programmes that build TPACK competence among medical educators, measuring impact on the quality of educator-designed interventions and on learner outcomes. Institutional investment in faculty capacity may yield greater returns than investment in technology itself.

Embed Ethical Governance and Participatory Design

AI-gamified medical education systems make consequential decisions about what trainees and patients learn and when. Future development should embed participatory design from the outset, with learners, educators, patients, ethicists, and regulators contributing to gamification mechanics and AI personalisation strategies.⁶⁰ Research priorities include developing ethical governance frameworks tailored to educational AI, evaluating algorithmic bias in adaptive learning systems across demographic groups, and establishing learner-autonomy safeguards that preserve agency within personalised pathways.

Explore Generative AI and Immersive Technologies for Next-Generation Learning

Generative AI and extended reality (XR) technologies are reshaping the design space for medical education. Large language models can dynamically generate clinical scenarios, virtual-patient dialogue, and literacy-matched patient education content; XR platforms enable immersive simulation for procedural training, rehabilitation, and exposure-based therapeutic education.⁶⁹ Future work should rigorously evaluate combinations of generative AI with reinforcement-learning-based adaptive instruction, addressing content fidelity, hallucination risk, and educational effectiveness across the Kirkpatrick hierarchy from reaction to results. Particular attention should be paid to whether immersive technologies improve transfer to clinical practice or merely enhance the in-simulation experience.

Synthesis

These seven priorities are interdependent. Theoretically grounded design (priority 1) is meaningless without robust outcome measurement (priority 2); validated learning analytics (priority 3) require privacy-preserving infrastructure (priority 4) to scale across institutions; and faculty capacity (priority 5), ethical governance (priority 6), and frontier technologies (priority 7) determine whether the field will mature into a discipline or remain a collection of pilots.

Figure 4 visualises this interdependence, showing the seven priorities organised into three mutually reinforcing clusters that converge on a mature, integrated discipline.

Conclusion

The convergence of artificial intelligence and gamification holds genuine promise for medical and health professional education. Together, they offer adaptive, engaging, and pedagogically rigorous learning experiences that neither technology can deliver alone — personalising content to the individual learner, sustaining motivation through well-designed game mechanics, and providing the real-time formative feedback that traditional instruction struggles to scale.

As this review demonstrates, however, the field remains constrained by weak theoretical grounding, inadequate outcome measurement, under-validated adaptive systems, algorithmic bias, equity gaps, and regulatory uncertainty. The operational definition, integration matrix, and ten real-world examples presented here offer a concrete vocabulary for designing and evaluating AI-enhanced gamified interventions in medical education, while the seven-priority research agenda translates aspiration into actionable next steps.

Progress will depend on interdisciplinary collaboration between learning scientists, clinicians, and AI engineers; rigorous active-comparator trials; transparent and reproducible algorithmic reporting; equitable deployment across diverse learner populations; and faculty capacity-building. If these conditions are met, AI-enhanced gamification can mature from a collection of pilots into a discipline that meaningfully advances how the next generation of clinicians and patients learn.

Acknowledgment

The authors confirm that Figures 1–4 were generated using ChatGPT Images 2.0, OpenAI, on 27 June 2026. The figures were generated from author-prepared prompts and were reviewed by the authors for scientific accuracy, relevance, and appropriateness for publication. No separate permission from OpenAI was required, as the images were generated by the authors using ChatGPT Images 2.0/OpenAI and, under the applicable OpenAI Terms of Use, the authors own the generated output to the extent permitted by law. The authors confirm that they reviewed the relevant terms of use, that the figures are suitable for publication, and that no third-party copyrighted images, patient-identifiable material, confidential data, or restricted content were used as inputs. The authors therefore confirm that they have the right to publish the AI-generated images in this manuscript, and that any permissions required have been obtained.

Disclosure

Dr. Md Anwarul Azim Majumder is the Editor-in-Chief of *Advances in Medical Education and Practice* (DovePress/Taylor & Francis) and a member of the Taylor & Francis Editorial Advisory Board. The other authors report no conflicts of interest in this work.

References

1. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? *Int J Educ Tech Higher Educ*. 2019;16(1):39. doi:10.1186/s41239-019-0171-0
2. Bond M, Bedenlier S, Marín VI, Händel M. Emergency remote teaching in higher education: mapping the first global online semester. *Int J Educ Technol High Educ*. 2021;18(1):50. doi:10.1186/s41239-021-00282-x
3. Organization WH. *Noncommunicable Diseases: Key Facts*. Geneva: World Health Organization;2023. Available from: <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases>. Accessed July 7, 2026.
4. Naghavi M, Ong KL, Aali A, et al. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2024;403(10440):2100–2132.
5. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, et al. Serious gaming and gamification education in health professions: systematic review. *J Med Internet Res*. 2019;21(3):e12994. doi:10.2196/12994
6. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7
7. Jayaraman P, Desman J, Sabounchi M, Nadkarni GN, Sakhuja A. A primer on reinforcement learning in medicine for clinicians. *NPJ Digit Med*. 2024;7(1):337. doi:10.1038/s41746-024-01316-0
8. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon DG. using game-design elements in non-gaming contexts. In: *CHI '11 Extended Abstracts on Human Factors in Computing Systems*. Vancouver, BC, Canada: Association for Computing Machinery; 2011:2425–2428.

9. Koivisto J, Hamari J. The rise of motivational information systems: a review of gamification research. *Int J Inform Manag*. 2019;45(C):191–210. doi:10.1016/j.ijinfomgt.2018.10.013
10. Jonassen DH. Objectivism versus constructivism: do we need a new philosophical paradigm? *Educ Technol Res Dev*. 1991;39(3):5–14. doi:10.1007/BF02296434
11. VanLehn K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ Psychol*. 2011;46(4):197–221. doi:10.1080/00461520.2011.611369
12. Edwards EA, Lumsden J, Rivas C, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open*. 2016;6(10):e012447. doi:10.1136/bmjopen-2016-012447
13. Blok AC, Sadasivam RS, Amante DJ, et al. Gamification to motivate the unmotivated smoker: the “take a break. Digital Health Intervention. *Games Health J*. 2019;8(4):275–284.
14. Tolks D, Schmidt JJ, Kuhn S. The role of AI in serious games and gamification for health: scoping review. *JMIR Serious Games*. 2024;12:e48258. doi:10.2196/48258
15. Damaševičius R, Maskeliūnas R, Blažauskas T. Serious games and gamification in healthcare: a meta-review. *Information*. 2023;14(2):105. doi:10.3390/info14020105
16. Mishra P, Koehler MJ. Technological pedagogical content knowledge: a framework for teacher knowledge. *Teachers College Record*. 2006;108(6):1017–1054. doi:10.1111/j.1467-9620.2006.00684.x
17. Mayer RE. Cognitive Theory of Multimedia Learning. In: Mayer RE, editor. *Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press; 2005:31–48.
18. Merrill MD. First principles of instruction. *Edu Technol Res Dev*. 2002;50(3):43–59. doi:10.1007/BF02505024
19. Werbach K, Hunter D. *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Digital Press; 2012.
20. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. *J Biomed Inform*. 2017;71:31–48. doi:10.1016/j.jbi.2017.05.011
21. Kirkwood A, Price L. Technology-enhanced learning and teaching in higher education: what is ‘enhanced’ and how do we know? A critical literature review. *Learning Media Technol*. 2014;39(1):6–36. doi:10.1080/17439884.2013.770404
22. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol*. 2000;55(1):68–78. doi:10.1037/0003-066X.55.1.68
23. Ng JY, Ntoumanis N, Thøgersen-Ntoumani C, et al. Self-determination theory applied to health contexts: a meta-analysis. *Perspect Psychol Sci*. 2012;7(4):325–340. doi:10.1177/1745691612447309
24. Csikszentmihalyi M. *Flow: The Psychology of Optimal Experience*. Harper & Row; 1990.
25. Hamari J, Koivisto J, Sarsa H. Does gamification work? A literature review of empirical studies on gamification. 47th Hawaii International Conference on System Sciences 2014. p. 3025–3034.
26. Y-k C. Actionable Gamification: beyond Points, Badges and Leaderboards: octalysis Media; 2016.
27. Mohanty S, Bagavan P. A bibliometric analysis of the use of the gamification octalysis framework in training. *Humanities Soc Sci Commun*. 2023;10:Article1. doi:10.1057/s41599-023-02243-3
28. Vygotsky LS. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press; 1978.
29. Siemens G. Connectivism: a learning theory for the digital age. *Int J Instruct Technol Distance Learn*. 2005;2(1):3–10.
30. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med*. 2013;46(1):81–95. doi:10.1007/s12160-013-9486-6
31. Wang YF, Hsu YF, Fang KT, Kuo LT. Gamification in medical education: identifying and prioritizing key elements through Delphi method. *Med Educ Online*. 2024;29(1):2302231. doi:10.1080/10872981.2024.2302231
32. Makani A. Onco-shikshak: an AI-native adaptive learning ecosystem for medical oncology education. *medRxiv*. 2026;2026:2026.02.23.26346944.
33. Evans KH, Daines W, Tsui J, Strehlow M, Maggio P, Shieh L. Septris: a novel, mobile, online, simulation game that improves sepsis recognition and management. *Acad Med*. 2015;90(2):180–184. doi:10.1097/ACM.0000000000000611
34. Boeker M, Andel P, Vach W, Frankenschmidt A. Game-based E-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. *PLoS One*. 2013;8(12):e82328. doi:10.1371/journal.pone.0082328
35. Ismail MA, Ahmad A, Mohammad JA, Fakri N, Nor MZM, MNM Pa. Using Kahoot! as a formative assessment tool in medical education: a phenomenological study. *BMC Med Educ*. 2019;19(1):230. doi:10.1186/s12909-019-1658-z
36. Ghafouri R, Namaghi SG, Khoshgouei B. The effect of gamification on the medication knowledge, performance and satisfaction of nurses in continued medical education: a quasi-experimental study. *PLoS One*. 2025;20(9):e0331372. doi:10.1371/journal.pone.0331372
37. Larsen CR, Oestergaard J, Ottesen BS, Soerensen JL. The efficacy of virtual reality simulation training in laparoscopy: a systematic review of randomized trials. *Acta Obstet Gynecol Scand*. 2012;91(9):1015–1028. doi:10.1111/j.1600-0412.2012.01482.x
38. Humm G, Mohan H, Fleming C, et al. The impact of virtual reality simulation training on operative performance in laparoscopic cholecystectomy: meta-analysis of randomized clinical trials. *BJS Open*. 2022;6(4):1.
39. He Y, Chen X, Ma H, Zhao R, Zhou H, Yang Y. Effectiveness of escape room in medical education: a systematic review and meta-analysis. *Games Health J*. 2024;13(6):419–427. doi:10.1089/g4h.2023.0070
40. Salimi S, Asgari Z, Mohammadnejad A, Teimazi A, Bakhtiari M. Efficacy of virtual reality and augmented reality in anatomy education: a systematic review and meta-analysis. *Anat Sci Educ*. 2024;17(9):1668–1685. doi:10.1002/ase.2501
41. Elzeky MEH, Elhabashy HMM, Ali WGM, Allam SME. Effect of gamified flipped classroom on improving nursing students’ skills competency and learning motivation: a randomized controlled trial. *BMC Nurs*. 2022;21(1):316. doi:10.1186/s12912-022-01096-6
42. Lee M, Shin S, Lee M, Hong E. Educational outcomes of digital serious games in nursing education: a systematic review and meta-analysis of randomized controlled trials. *BMC Med Educ*. 2024;24(1):1458. doi:10.1186/s12909-024-06464-1
43. Brons A, Wang S, Visser B, Kröse B, Bakkes S, Velkamp R. Machine learning methods to personalize persuasive strategies in mhealth interventions that promote physical activity: scoping review and categorization overview. *J Med Internet Res*. 2024;26:e47774. doi:10.2196/47774
44. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. 2nd ed. MIT Press; 2018.
45. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc*. 2018;25(9):1248–1258. doi:10.1093/jamia/ocy072

46. Baker RS, Inventado PS. Educational Data Mining and Learning Analytics. In: Larusson JA, White B, editors. *Learning Analytics: From Research to Practice*. New York: Springer New York; 2014:61–75.
47. Varga G, Stoicu-Tivadar L, Nicola S. Serious gaming and artificial intelligence in rehabilitation of rheumatoid arthritis. *Stud Health Technol Inform*. 2022;295:562–565. doi:10.3233/SHIT220790
48. Aguilera A, Arévalo Avalos M, Xu J, et al. Effectiveness of a digital health intervention leveraging reinforcement learning: results from the diabetes and mental health adaptive notification tracking and evaluation (DIAMANTE) randomized clinical trial. *J Med Internet Res*. 2024;26:e60834. doi:10.2196/60834
49. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAI) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med*. 2018;52(6):446–462. doi:10.1007/s12160-016-9830-8
50. Deliu N, Williams J, Chakraborty B. Reinforcement learning in modern biostatistics: constructing optimal adaptive interventions. *Int Stat Rev*. 2024;92(2):e2470.
51. Vandelandotte C, Trost S, Hodgetts D, et al. Increasing physical activity using an just-in-time adaptive digital assistant supported by machine learning: a novel approach for hyper-personalised mHealth interventions. *J Biomed Inform*. 2023;144:104435. doi:10.1016/j.jbi.2023.104435
52. Cook DA, Overgaard J, Pankratz VS, Del Fiol G, Aakre CA. Virtual patients using large language models: scalable, contextualized simulation of clinician-patient dialogue with feedback. *J Med Internet Res*. 2025;27:e68486. doi:10.2196/68486
53. VanLehn K. The behavior of tutoring systems. *Int J Artificial Intell Educ*. 2006;16(3):227–265. doi:10.3233/IRG-2006-16(3)02
54. Wang C, Lee C, Shin H. Digital therapeutics from bench to bedside. *NPJ Digit Med*. 2023;6(1):38. doi:10.1038/s41746-023-00777-z
55. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–444. doi:10.1038/nature14539
56. Shute VJ. Stealth assessment in computer-based games to support learning. In: Tobias S, Fletcher JD, editors. *Computer Games and Instruction: Information Age*. 2011:503–524.
57. Nishi SK, Kavanagh ME, Ramboanga K, et al. Effect of digital health applications with or without gamification on physical activity and cardiometabolic risk factors: a systematic review and meta-analysis of randomized controlled trials. *EClinicalMedicine*. 2024;76:102798. doi:10.1016/j.eclim.2024.102798
58. Klasnja P, Hekler EB, Shiffman S, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. *Health Psychol*. 2015;34s:1220–1228. doi:10.1037/hea0000305
59. Eysenbach G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res*. 2011;13(4):e126. doi:10.2196/jmir.1923
60. Cohen IG, Mello MM. Big data, big tech, and protecting patient privacy. *JAMA*. 2019;322(12):1141–1142. doi:10.1001/jama.2019.11365
61. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. *J Med Internet Res*. 2020;22(6):e19361. doi:10.2196/19361
62. Denoni-Buján M, Marcen C, Gracia-Gil A, Casanovas López R, Coral-Aguilar S. Challenges of innovation through gamification in the classroom. *Educ Sci*. 2025;15(10):1341. doi:10.3390/educsci15101341
63. Food US, Administration D. *Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions*. Silver Spring: FDA; 2023.
64. Carlier S, De Backere F, De Turck F. Personalised serious games and gamification in healthcare: survey and future research directions. *ACM Computing Surveys*. 2024;56(8):1–34.
65. Schmidt L, Pawlitzki M, Renard BY, Meuth SG, Masannek L. The three-year evolution of Germany's Digital Therapeutics reimbursement program and its path forward. *NPJ Digit Med*. 2024;7(1):139. doi:10.1038/s41746-024-01137-1
66. Abbas SR, Abbas Z, Zahir A, Lee SW. Federated Learning in Smart Healthcare: a Comprehensive Review on Privacy, Security, and Predictive Analytics with IoT Integration. *Healthcare*. 2024;12(24):2587. doi:10.3390/healthcare12242587
67. Wang M, Zhu J, Qian W, Gu H. Integrating artificial intelligence and gamification in rehabilitation: a scoping review. *Entertainment Computing*. 2025;55:101011. doi:10.1016/j.entcom.2025.101011
68. Siemens G, Long P. Penetrating the fog: analytics in learning and education. *EDUCAUSE Rev*. 2011;46(5):30–40.
69. Herur-Raman A, Almeida ND, Greenleaf W, Williams D, Karshenas A, Sherman JH. Next-generation simulation—integrating extended reality technology into medical education. *Front Virtual Reality*. 2021. doi:10.3389/frvir.2021.693399

Advances in Medical Education and Practice

Publish your work in this journal

Advances in Medical Education and Practice is an international, peer-reviewed, open access journal that aims to present and publish research on Medical Education covering medical, dental, nursing and allied health care professional education. The journal covers undergraduate education, postgraduate training and continuing medical education including emerging trends and innovative models linking education, research, and health care services. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/advances-in-medical-education-and-practice-journal>

Dovepress
Taylor & Francis Group