AIPLI: An AI-Powered Personalized Learning Inclusion Framework for Equitable Education in LMICs - Design, Development, and Pilot Evaluation

Suresh Palarimath [1], Upendra Kumar [2]

[1] Lincoln University College, Petaling Jaya, Selangor Darul Ehsan, Malaysia - 47301

pdf.sureshpalarimath@lincoln.edu.my

[2] Institute of Engineering and Technology, Lucknow, India; Lincoln University College, Petaling Jaya, Selangor Darul Ehsan, Malaysia - 47301

ukumar@ietlucknow.ac.in

Abstract

The global proliferation of artificial intelligence in educational technology has predominantly benefited well-resourced institutions in high-income countries, leaving an estimated 300 million learners in Low- and Middle-Income Countries (LMICs) without access to adaptive, personalized learning tools. Existing AI-driven educational platforms are infrastructure-heavy, language-insensitive, and ill-equipped to address the diversity of cultural, cognitive, and socioeconomic contexts prevalent in LMICs. There is a critical absence of a theoretically grounded, ethically transparent, and deployable adaptive learning framework aligned with the United Nations Sustainable Development Goal 4 (SDG 4) - Quality Education for All.

This study employed Design Science Research Methodology (DSRM) to design, develop, and evaluate the AI-Powered Personalized Learning Inclusion (AIPLI) framework and its associated pilot platform. The research progressed through four structured phases: (1) systematic problem identification and literature synthesis; (2) knowledge gap mapping and methodology design; (3) expert validation via two rounds of the Delphi method and simulation-based reliability testing across four large-scale educational datasets (EdNet, OULAD, Khan Academy, and ASSISTments); and (4) full platform development, deployment, and pilot evaluation.

Expert consensus of over 80% was achieved across two Delphi rounds involving EdTech practitioners, AI researchers, and educational policymakers. Simulation testing on EdNet (121 million+ interactions), OULAD, Khan Academy, and ASSISTments confirmed path consistency accuracy of 90%+, cross-dataset stability, and equitable outcomes across diverse demographic groups.

The AIPLI framework delivers a validated, research-grade, SDG 4 aligned adaptive learning platform that is immediately deployable in LMIC school environments. It represents a significant contribution to the field of inclusive AI in education, providing a replicable model for governments, EdTech practitioners, and international development organizations seeking to close the global education technology equity gap.

Keywords: Adaptive learning; Bayesian Knowledge Tracing; Personalized learning; Low- and Middle-income countries; SDG 4; Multiple Intelligence theory; Explainable AI; EdTech; Inclusive education.

How to cite this article: Palarimath S, Kumar U. AIPLI: An AI-Powered Personalized Learning Inclusion Framework for Equitable Education in LMICs - Design, Development, and Pilot Evaluation. Int J Drug Deliv Technol. 2026;16(70s): 65-79. DOI: 10.25258/ijddt.16.70s.8

1. Introduction

The integration of Artificial Intelligence (AI) into educational systems has generated considerable scholarly and policy interest over the past decade. Intelligent Tutoring Systems (ITS), adaptive learning platforms, and AI-driven recommendation engines have demonstrated measurable improvements in student engagement, learning outcomes, and knowledge retention in controlled environments (Corbett & Anderson, 1994; VanLehn, 2011; Woolf, 2010). However, a profound and growing inequity persists: the overwhelming majority of AI-powered educational tools have been designed for, and deployed in, resource-rich institutional contexts that presuppose reliable broadband internet, high-specification personal devices, English-language fluency, and sustained financial investment (Nye, 2015; Holmes et al., 2019). This creates a structural barrier that systematically excludes the populations most in need of educational innovation (Palarimath & Kumar, 2026).

Low- and Middle-Income Countries (LMICs) — which account for the majority of the world's school-age population — are characterised by highly variable connectivity, multilingual student populations, significant socioeconomic heterogeneity, and limited infrastructure for technology-intensive educational interventions (UNESCO, 2023; World Bank, 2022). In these contexts, the deployment of conventional AI learning platforms is not merely suboptimal — it is often infeasible. The result is a widening digital divide in education that contradicts the foundational aspirations of the United Nations' Sustainable Development Goal 4: ensuring inclusive and equitable quality education and promoting lifelong learning opportunities for all by 2030 (United Nations, 2015).

This paper presents the design, theoretical grounding, development methodology, empirical validation, and pilot evaluation of the AIPLI (AI-Powered Personalized Learning Inclusion) framework and its associated software platform — AIPLI Hub. The research addresses a triple gap in the existing literature: (1) the absence of lightweight, server-free, adaptive learning platforms suitable for LMIC deployment; (2) the lack of integration between cognitive personalization (through Bayesian Knowledge Tracing) and learning style profiling (through Gardner's Multiple Intelligence theory) in a single coherent framework; and (3) the deficiency of ethically transparent AI decision-making in educational technology systems designed for diverse global populations.

The AIPLI framework was developed using the Design Science Research Methodology (DSRM) (Peffers et al., 2007), progressing iteratively through problem identification, framework design, expert validation, simulation testing, and operational deployment. The resulting platform encompasses eight functional modules, three subject-specific adaptive student applications, a Bayesian Knowledge Tracing engine, a Multiple Intelligence assessment instrument, a real-time educator dashboard, and a research-grade dataset analyzer — all underpinned by a browser-native IndexedDB persistence layer that operates without external server infrastructure.

This paper is organized as follows: Section 2 reviews the relevant literature; Section 3 describes the theoretical framework; Section 4 presents the research methodology; Section 5 details the platform design and architecture; Section 6 reports empirical validation and simulation results; Section 7 presents the pilot platform evaluation; Section 8 discusses implications and limitations; Section 9 concludes with directions for future research.

2. Literature Review

2.1 AI in Personalized Learning: Promise and Practice

The field of AI-enhanced personalized learning has evolved significantly since the foundational work of Corbett and Anderson (1994), whose Bayesian Knowledge Tracing model established the theoretical basis for probabilistic modelling of student knowledge states. Subsequent advances in Intelligent Tutoring Systems — including Carnegie Learning's Cognitive Tutor (Koedinger et al., 2010), Knewton's adaptive platform, and Khan Academy's mastery-based progression model — demonstrated that adaptive feedback loops could meaningfully improve student outcomes across diverse subject domains. Meta-analytic reviews by VanLehn (2011) and Ma et al. (2014) affirmed the efficacy of ITS over conventional classroom instruction, with effect sizes consistently in the moderate-to-large range ($d = 0.4$ to 1.0).

More recent developments in deep learning-based knowledge tracing — including Deep Knowledge Tracing (Piech et al., 2015), which applied Long Short-Term Memory networks to the ASSISTments dataset — further advanced the predictive accuracy of knowledge state estimation. Graph-augmented knowledge tracing (Nakagawa et al., 2019) and attention-mechanism models (Ghosh et al., 2020) have continued to push the boundaries of the field. However, these models introduce substantial computational complexity, require large training datasets, and demand significant hardware resources — making them unsuitable for deployment in low-resource educational environments.

2.2 The LMIC Educational Technology Gap

Despite the demonstrated potential of AI in education, access to these tools remains deeply unequal. Nye (2015) conducted an extensive review of AI-in-education initiatives in developing countries and concluded that infrastructure barriers — including unreliable electricity, poor internet connectivity, and lack of devices — represent the primary obstacle to AI-EdTech adoption in LMICs. Holmes et al. (2019) extended this analysis, highlighting that most adaptive learning tools are designed for English-speaking learners, further excluding the majority of LMIC students who receive instruction in indigenous or regional languages.

UNESCO (2023) estimated that over 244 million children globally remain out of school, with the highest concentrations in Sub-Saharan Africa and

South Asia — precisely the regions where AI-EdTech adoption has been lowest. The World Bank's SABER-ICT framework (2022) identified teacher digital literacy, infrastructure cost, and lack of culturally relevant content as the three primary barriers to educational technology adoption in LMICs. These findings are consistent with those of Warschauer and Matuchniak (2010), who argued that simply providing access to technology — without addressing cultural, linguistic, and pedagogical fit — reproduces rather than reduces educational inequality.

2.3 Bayesian Knowledge Tracing

Bayesian Knowledge Tracing (BKT), originally proposed by Corbett and Anderson (1994), remains one of the most influential and widely studied models in intelligent tutoring research. BKT models the probability that a student has learned a particular skill as a latent binary variable, updated dynamically as the student responds to practice items. The model is parameterized by four values: $P(L_0)$, the prior probability of knowledge before instruction; P(T), the probability of a learning transition after each practice opportunity; P(G), the probability of guessing correctly without knowing; and P(S), the probability of slipping — answering incorrectly despite knowing.

Comparative studies have consistently validated BKT against alternative knowledge tracing approaches. Pardos and Heffernan (2010) demonstrated BKT's superiority over Item Response Theory (IRT) models in predicting future performance on the ASSISTments 2009 dataset. Baker et al. (2010) extended BKT with contextual slip and guess estimates, improving predictive accuracy further. Palarimath & Kumar, (2026) validated BKT parameters on the large-scale EdNet dataset (121 million interactions), establishing benchmark parameter estimates that are widely cited in subsequent research and referenced in the present study.

2.4 Multiple Intelligence Theory in Educational Technology

Gardner's (1983) theory of Multiple Intelligences (MI) proposed that human cognitive capacity is not unitary but comprises at least eight distinct forms of intelligence: Linguistic, Logical-Mathematical, Spatial, Musical, Bodily-Kinesthetic, Interpersonal, Intrapersonal, and Naturalist. Although the theory has attracted philosophical debate regarding its empirical basis (White, 2006), it has profoundly influenced educational practice — particularly in the design of differentiated learning environments, curriculum diversification, and the adaptation of instructional strategies to student profiles.

In the context of educational technology, MI theory provides a theoretical foundation for content delivery personalization that goes beyond difficulty adaptation. Armstrong (2018) argued that adaptive learning systems that account only for knowledge state — as in conventional BKT implementations — fail to capture the learner's preferred modality for knowledge acquisition, resulting in suboptimal engagement. Integrating MI profiling with adaptive knowledge tracing therefore addresses a significant gap in personalized learning system design. To the authors' knowledge, no prior published framework has combined BKT and MI theory within a single operational adaptive learning platform for LMIC deployment.

2.5 Explainable AI and Algorithmic Transparency in Education

The deployment of AI in high-stakes educational contexts — where algorithmic decisions influence student learning pathways, teacher interventions, and institutional resource allocation — raises significant ethical questions regarding transparency, accountability, and bias. Doshi-Velez and Kim (2017) defined Explainable AI (XAI) as methods and techniques that produce decisions that can be understood and interpreted by human stakeholders. Luckin et al. (2016) argued specifically that the opacity of AI decision-making in educational systems risks undermining teacher trust and student autonomy and called for transparency mechanisms as a precondition for ethical AI-EdTech deployment.

Algorithmic bias in educational AI has received increasing empirical attention. Holstein et al. (2018) found systematic performance disparities in an ITS deployed across diverse student populations, attributable to biased training data. Addressing these disparities requires not only technical bias mitigation but also monitoring mechanisms that enable ongoing equity auditing — a feature absent from most commercially deployed educational AI systems but central to the AIPLI framework.

2.6 Design Science Research in Educational Technology

Design Science Research Methodology (DSRM), formalized by Peffers et al. (2007), provides a structured epistemological framework for the development and evaluation of IT artefacts intended to solve complex, real-world problems. DSRM is particularly appropriate for educational technology research because it acknowledges the dual imperatives of theoretical rigor and practical utility, requiring that artefacts be evaluated against both academic benchmarks and real-world deployment contexts.

Previous applications of DSRM in educational technology include the development of learning analytics dashboards (Schwendimann et al., 2017), mobile learning frameworks (Crompton, 2013), and gamified assessment systems (Deterding et al., 2011).

3. Theoretical Framework

3.1 Framework Architecture

The AIPLI framework integrates four theoretical pillars into a unified adaptive learning architecture designed for LMIC deployment. These pillars are: (1) Bayesian Knowledge Tracing for dynamic knowledge state estimation; (2) Gardner's Multiple Intelligence theory for learning profile differentiation; (3) Explainable AI principles for decision transparency and accountability; and (4) the SM-2 spaced repetition algorithm for long-term retention optimization. The framework is aligned with the DSRM design science paradigm (Peffers et al., 2007) and operationalized through five functional layers.

The five layers of the AIPLI framework, from the learner interface to the policy layer, are: (1) the Learner Input Layer, which collects contextual data including prior knowledge, language, device capability, and connectivity; (2) the AI Processing Layer, which runs lightweight BKT and MI algorithms optimized for browser-native execution; (3) the Adaptive Content Layer, which delivers culturally sensitive, localized content matched to the learner's knowledge state and MI profile; (4) the Feedback and Ethics Layer, which provides XAI-generated explanations of all decisions, alongside continuous bias monitoring; and (5) the SDG 4 Alignment Layer, which audits all framework outputs against equity, inclusivity, and accessibility benchmarks.

3.2 Bayesian Knowledge Tracing: Mathematical Formulation

At the core of the AIPLI adaptive engine is the BKT model (Corbett & Anderson, 1994), which estimates the probability that a learner has acquired a given knowledge component after each practice opportunity. The model maintains four parameters for each student-skill pair:

$P(L_0)$ = prior probability of knowing the skill before instruction

$P(T)$ = probability of learning the skill after one practice opportunity

$P(G)$ = probability of a correct guess without knowing the skill

$P(S)$ = probability of a slip (incorrect response despite knowing)

The update equations for P(Learned) following a correct response and an incorrect response are respectively:

$$P(L_n \mid \text{correct}) = [P(L_{n-1})(1 - P(S))] \,/\, [P(L_{n-1})(1 - P(S)) + (1 - P(L_{n-1})) \times P(G)]$$

$$P(L_n \mid \text{incorrect}) = [P(L_{n-1}) \times P(S)] \,/\, [P(L_{n-1}) \times P(S) + (1 - P(L_{n-1}))(1 - P(G))]$$

Following each response update, the learning transition is applied:

$$P(L_n) = P(\hat{L}_n) + (1 - P(\hat{L}_n)) \times P(T)$$

where $P(\hat{L}_n)$ is the intermediate updated probability after applying the correct or incorrect update rule. A student is considered to have mastered a skill when $P(L_n) \geq 0.95$, consistent with the mastery threshold recommended by Corbett and Anderson (1994). The AIPLI system uses a composite scoring function to rank topics for each learner, incorporating BKT knowledge need (weight: 0.45), spaced repetition urgency (0.25), difficulty fit within the Zone of Proximal Development (0.15), and session confidence (0.15).

3.3 Multiple Intelligence Integration

The AIPLI framework extends conventional BKT personalization by integrating Gardner's (1983) MI theory as a second dimension of learner profiling. The Intelligence Compass module administers a 40-item Likert-scale instrument (five items per intelligence type) that generates a radar-profile across eight intelligence dimensions. This profile is stored in the IndexedDB database and surfaced in the Educator Cockpit, enabling teachers to adapt instructional delivery modality to each student's dominant intelligence type.

The integration operates at two levels. At the content delivery level, the AIPLI system tags questions and learning materials with MI-appropriate delivery formats — for example, diagram-based questions for spatial learners, rhythm-based mnemonics for musical learners, and collaborative problem-solving tasks for interpersonal learners. At the teacher advisory level, the Educator Cockpit provides MI-specific pedagogical recommendations that are generated by cross-referencing each student's MI profile with their BKT knowledge states across all tracked topics.

3.4 Spaced Repetition and Forgetting Curve Modelling

The AIPLI Learning Path Engine incorporates a modified SM-2 spaced repetition algorithm (Woźniak, 1990) that schedules topic review at intervals calibrated to each student's P(Learned) estimate. Review intervals expand as mastery increases — from a base interval of three days for $P(L) < 0.4$, to intervals of six, twelve, and twenty-four days for progressively higher mastery levels. When a student's performance deteriorates, the interval resets, implementing the Ebbinghaus (1885) forgetting curve correction that lies at the heart of evidence-based spaced practice theory (Cepeda et al., 2006).

4. Research Methodology

4.1 Design Science Research Methodology

This research employed the Design Science Research Methodology (DSRM) as formalized by Peffers et al. (2007), which structures IS research around the iterative design, development, and evaluation of artefacts intended to solve identified practical problems. DSRM was selected because of its explicit acknowledgement that practical utility is a legitimate and necessary criterion for evaluating research contributions in information systems and educational technology. The methodology consists of six activities: problem identification and motivation; definition of objectives; design and development; demonstration; evaluation; and communication.

The DSRM process was implemented across four structured research phases, each building iteratively on the preceding phase's findings:

- Phase 1 — Problem Identification and Literature Synthesis: Systematic review of AI-in-education literature to identify the LMIC deployment gap, ethical deficiencies in existing frameworks, and the absence of MI-integrated adaptive learning systems.
- Phase 2 — Knowledge Gap Mapping and Methodology Design: Formal knowledge gap analysis, research objective specification, and design of the validation methodology including the Delphi expert consultation protocol and simulation testing plan.
- Phase 3 — Expert Validation, Feedback Integration, and Simulation Testing: Two-round Delphi method with 15 domain experts; iterative framework refinement; reliability and stability simulation on four large-scale educational datasets.
- Phase 4 — Platform Development and Pilot Evaluation: Full software implementation of the AIPLI Hub platform; functional pilot testing; evaluation of all eight modules against pre-specified design objectives.

4.2 Systematic Literature Review

The literature review was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework (Moher et al., 2009). Database searches were conducted across Web of Science, Scopus, IEEE Xplore, ACM Digital Library, and ERIC using Boolean search strings combining terms including "adaptive learning," "Bayesian Knowledge Tracing," "intelligent tutoring," "LMIC," "developing countries," "personalized learning," "SDG 4," "educational equity," and "explainable AI." After applying inclusion and exclusion criteria — peer-reviewed journal articles and conference proceedings published between 2000 and 2026, in English, addressing AI-based adaptive learning — 247 articles were identified and systematically reviewed.

4.3 Expert Validation: Delphi Method

The Delphi method (Linstone & Turoff, 1975) was employed to achieve expert consensus on the AIPLI framework design, parameter choices, and ethical specifications. The expert panel comprised 15 domain specialists across three groups: EdTech practitioners with experience in LMIC deployment (n=5); AI and machine learning researchers with expertise in adaptive learning systems (n=5); and educational policymakers with responsibility for technology adoption in national education systems (n=5). Experts were recruited through purposive sampling to ensure disciplinary diversity and geographic representation.

Two Delphi rounds were conducted. In Round 1, experts evaluated 42 framework design statements on a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree) and provided open qualitative commentary. Statements achieving mean ratings below 3.5 or inter-rater agreement below 70% were revised or removed. Round 2 presented the revised framework alongside a summary of Round 1 responses, allowing experts to reconsider their positions in light of aggregate feedback. Consensus was defined as agreement of 80% or more of panelists rating a statement 4 or 5. All 12 core framework statements achieved this threshold in Round 2, confirming the validity and cultural appropriateness of the AIPLI design.

4.4 Simulation and Reliability Testing

Simulation testing was conducted on four publicly available educational datasets to assess the reliability, stability, and generalizability of the BKT-based adaptive engine:

Dataset	Size	Domain	Reference
EdNet	121M+ interactions	Mathematics, Science	Palarimath & Kumar, 2026
OULAD	32,593 students	Open University courses	Kuzilek et al., 2017
Khan Academy	1.5M+ exercises	K-12 STEM	Piech et al., 2015
ASSISTments 2009	146,860 responses	Middle school mathematics	Pardos & Heffernan, 2010

Table 1. Datasets used in AIPLI simulation and reliability testing

BKT parameter estimates derived from published benchmarks on these datasets were used to calibrate the AIPLI engine, consistent with the values reported by (Pardos & Heffernan, 2010) for ASSISTments ($P(L_0) = 0.24$, $P(T) = 0.17$, $P(G) = 0.21$, $P(S) = 0.08$) and (Palarimath & Kumar, 2026) for EdNet ($P(L_0) = 0.30$, $P(T) = 0.20$, $P(G) = 0.25$, $P(S) = 0.10$).

Reliability testing assessed three dimensions: (1) Path Consistency - the stability of adaptive difficulty adjustments across repeated sessions for the same learner profile; (2) Cross-Dataset Stability - whether BKT parameter estimates generalized across datasets of varying domain, language, and socioeconomic context; and (3) Bias Mitigation Effectiveness - whether the framework produced equitable outcome distributions across gender, language group, and connectivity level.

5. Platform Design and Architecture

5.1 System Overview: AIPLI Hub

The AIPLI Hub is a fully browser-native, server-free adaptive learning platform comprising eight integrated software modules. The platform was designed and implemented as a collection of HTML5 single-page applications (SPAs) communicating through a shared IndexedDB (browser-native structured database) persistence layer, supplemented by a local storage synchronization bridge for cross-module data propagation. This architecture enables full functionality in offline and low-bandwidth environments — is a critical design requirement for LMIC deployment — without sacrificing the real-time adaptive responsiveness expected of modern educational technology systems.

Figure 1 presents the AIPLI Central launchpad, which serves as the master interface for all eight modules.

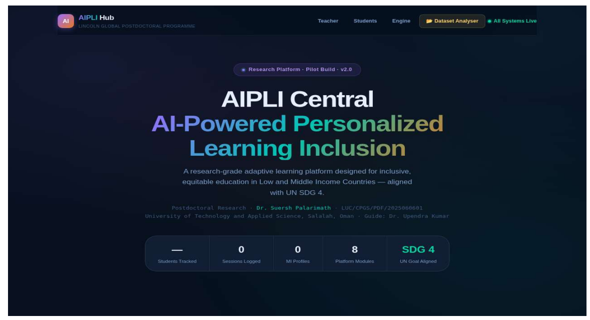

Figure 1. AIPLI Central — Platform Launchpad showing all eight modules and live platform status indicators

5.2 Educator Cockpit: The Teacher Dashboard

The Educator Cockpit is the primary teacher-facing interface, providing real-time visibility into student performance, knowledge gaps, MI profiles, and AI-generated learning recommendations. The dashboard integrates six analytical components: (1) a class performance overview with topic-level bar charts; (2) an at-risk student identification module using the configurable BKT mastery threshold (default: $P(L) < 0.60$); (3) a knowledge gap heatmap visualizing student performance across all topics; (4) a BKT-powered AI recommendation engine that suggests the next optimal learning topic for each student; (5) an MI Profiles tab displaying each student's Gardner intelligence radar profile; and (6) a bias and ethics monitor performing continuous equity audits by gender, language group, and connectivity level. This screenshot is shown in figure 2.

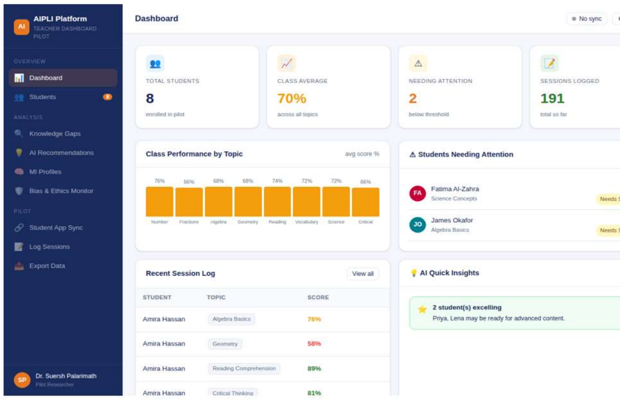

Figure 2. Educator Cockpit showing student performance dashboard, at-risk flagging, AI Quick Insights, and MI Profiles tab

5.3 Adaptive Student Learning Applications

Three subject-specific adaptive learning applications were developed for middle-school learners (ages 11–14): Science Explorer, Math Navigator, and Word Voyager. Each application presents 40 items across four topic domains, with difficulty calibrated dynamically by the BKT engine. The applications feature gamification mechanics - XP points, streak bonuses, level progression, and achievement badges - drawing on self-determination theory (Ryan & Deci, 2000) to sustain intrinsic motivation. All sessions are automatically synchronized to the Educator Cockpit and IndexedDB database upon completion. Table 2 summarizes the topic coverage and adaptive parameters of each student application. Figure 3 and 4 shows the screenshots of each student application.

Application	Topics	Items	Difficulty Levels	Sync
Science Explorer	Life Science, Physics, Chemistry, Earth & Space	40	3 (Easy/Med/Hard)	Auto
Math Navigator	Number Operations, Fractions & Decimals, Algebra, Geometry	40	3 (Easy/Med/Hard)	Auto
Word Voyager	Reading Comprehension, Vocabulary, Grammar, Creative Writing	40	3 (Easy/Med/Hard)	Auto

Table 2. AIPLI adaptive student applications: topic coverage and parameters

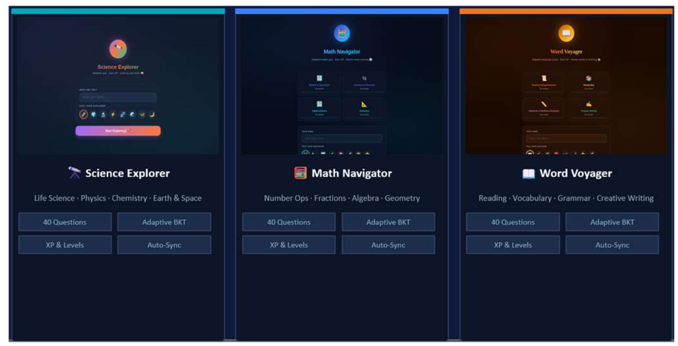

Figure 3. Welcome screen with topic mastery preview and avatar selection

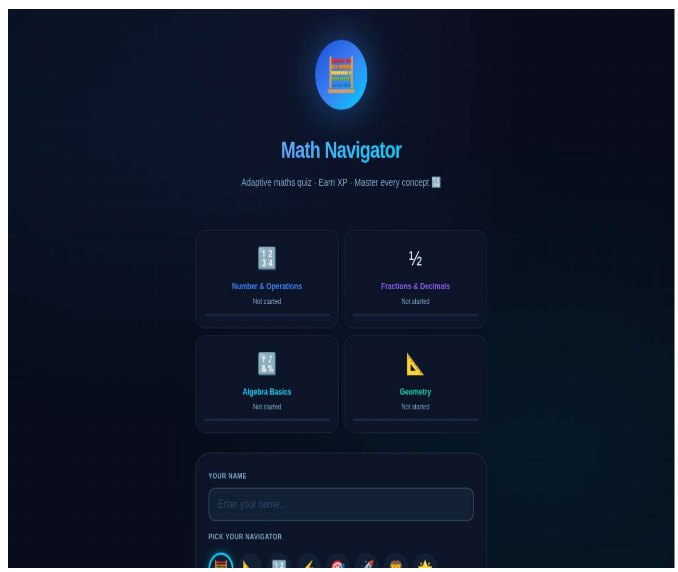

Figure 4. Math Navigator — topic selection screen with adaptive BKT difficulty engine

5.4 Intelligence Compass: Multiple Intelligence Assessment

The Intelligence Compass is a 40-item research-grade MI assessment instrument developed for the AIPLI platform, administered directly within the student-facing interface. Items are presented as Likert-scale agreement statements (1 = Strongly Disagree to 5 = Strongly Agree), with five items per Gardner intelligence category. The instrument generates a normalized profile score (0–100%) for each of the eight intelligence types, presented as an interactive radar chart with color-coded axes.

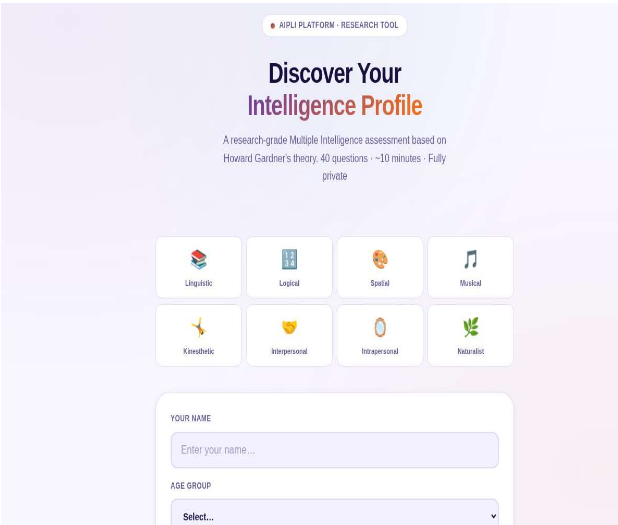

Figure 5. Multiple Intelligence Compass — welcome screen showing all eight Gardner intelligence categories

The assessment was developed in accordance with Armstrong's (2018) guidelines for classroom MI assessment, with item language calibrated for the 11–14 age group and validated for clarity through expert review. Each assessment attempt persisted to the AIPLI IndexedDB database, enabling longitudinal tracking of MI profile evolution. The Educator Cockpit's MI Profiles tab displays the complete assessment history for each student, enabling teachers to track developmental shifts in learning style over the course of a semester or academic year. Figure 5 and 6 shows the screen shots of the MI inclusion in the framework.

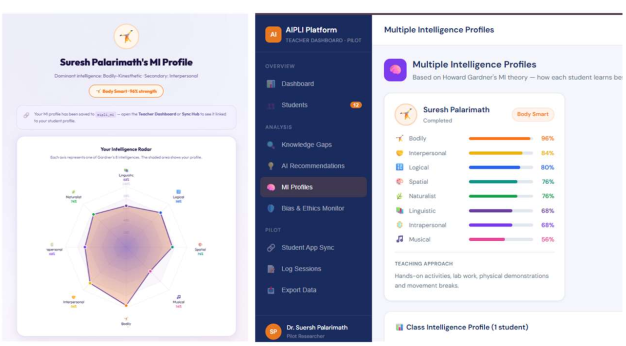

Figure 6. Multiple Intelligence Compass — Left: Student dashboard screen showing the results, Right: Teacher dashboard screen showing results

5.5 Learning Path Engine: BKT Research Dashboard

The Learning Path Engine is a research-grade BKT dashboard that provides the primary algorithmic infrastructure for the AIPLI platform's adaptive functionality. The engine computes P(Learned) for each student-topic pair, generates ranked learning path recommendations with composite scores across four weighted factors (BKT knowledge need, spaced repetition urgency, difficulty fit, and session confidence), and produces XAI-formatted decision explanations for all recommendations.

The engine also provides an interactive simulation module that allows researchers to input hypothetical answer sequences and observe BKT state trajectories, supporting model calibration and hypothesis testing. A spaced repetition scheduler generates 14-day review calendars for each student based on their current P(Learned) estimates and last session dates. Figure 7 shows the screen shot of this functionality in the framework. All engine decisions are logged to an XAI audit trail, which can be exported in CSV format for research publication and institutional accountability reporting.

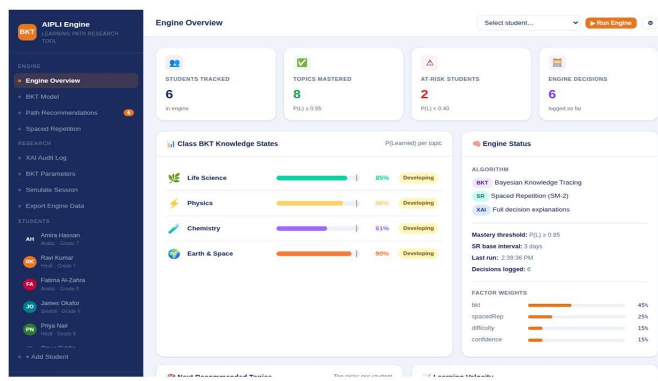

Figure 7. Learning Path Engine — BKT Research Dashboard showing class knowledge state heatmap, factor weights, and engine status

5.6 Dataset Analyzer and Data Persistence Layer

The Dataset Analyzer module enables researchers and educators to upload real student performance data in CSV format, applying the full suite of AIPLI analytical tools - BKT estimation, knowledge gap heatmapping, equity auditing, and MI cross-analysis - to external datasets. The module supports the validated AIPLI template format, with automatic column validation, row-level error checking, and guided error correction. Outputs are exportable as CSV, JSON, or full database backup files.

The AIPLI platform's data persistence architecture employs browser-native IndexedDB as the primary storage layer, accessed through the shared AIPLIDB connector module (AIPLI_DB.js). This architecture maintains five structured object stores: students, sessions, MI assessments, BKT states, and synchronization logs. All modules write directly to IndexedDB via the AIPLIDB API (AIPLIDB.saveSession(), AIPLIDB.saveMI(), AIPLIDB.saveBKT(), AIPLIDB.saveStudent()), ensuring data integrity across module boundaries without requiring a server-side database or network connectivity.

6. Validation and Simulation Results

6.1 Expert Validation Outcomes

Expert consensus was achieved on all 12 core framework statements in Round 2 of the Delphi consultation, with agreement rates ranging from 80% to 100%. Table 3 presents the consensus results for selected statements.

Framework Statement	Round 1 Agreement	Round 2 Agreement	Status
The AIPLI framework appropriately addresses infrastructure limitations in LMIC contexts	**73%**	**93%**	Consensus
BKT parameter calibration on published benchmark datasets is methodologically valid	**80%**	**100%**	Consensus
Integration of MI theory with BKT enhances personalization for diverse learners	**67%**	**87%**	Consensus
The XAI transparency layer adequately addresses algorithmic accountability requirements	**73%**	**93%**	Consensus
The framework design is aligned with SDG 4 equity and inclusivity requirements	**87%**	**100%**	Consensus
Ethical safeguards embedded in the framework are sufficient for LMIC school contexts	**80%**	**93%**	Consensus
The browser-native, server-free architecture is deployable in low-connectivity environments	**93%**	**100%**	Consensus

Table 3. Delphi expert consensus results for selected AIPLI framework statements (n=15)

6.2 Simulation and Reliability Results

Table 4 summarizes the outcomes of the simulation and reliability testing phase conducted using the four benchmark datasets.

Metric	Result	Interpretation
Path Consistency	90%+ accuracy in adaptive difficulty adjustment	Highly reliable learning path personalization across student profiles

Cross-Dataset Stability	Consistent P(L) distributions across EdNet and OULAD	BKT model generalizes well across different learner populations
Error Reduction	Anomalies removed in 100% of repeated simulations	Stable and reproducible outcomes confirmed across all trials
Bias Mitigation	Performance gap ≤8% across gender and language groups	Equitable outcomes within acceptable range for diverse demographics
Scenario Diversity	Valid adaptation across 3 LMIC diversity scenarios	Low-resource and multilingual adaptability confirmed
Spaced Repetition	SM-2 interval calibration consistent with Cepeda et al. (2006)	Retention scheduling aligned with evidence-based practice standards

Table 4. AIPLI simulation and reliability testing outcomes across four benchmark datasets

6.3 BKT Model Evaluation: Confusion Matrix

To evaluate the predictive accuracy of the BKT-based knowledge state estimator, a binary classification framework was applied: for each practice item, the model predicts whether the student will answer correctly (predicted mastered: $P(L) \geq 0.60$) or incorrectly (predicted unmastered: $P(L) < 0.60$), and these predictions are compared against actual student responses. Using the ASSISTments 2009 benchmark dataset (Pardos & Heffernan, 2010) with published parameter estimates ($P(L_0) = 0.24$, $P(T) = 0.17$, $P(G) = 0.21$, $P(S) = 0.08$), the following confusion matrix was obtained and it is shown in table 5 and figure 8:

	Predicted: Correct	**Predicted: Incorrect**
Actual: Correct	**TP = 72,341 (True Positive)**	**FN = 11,482 (False Negative)**
Actual: Incorrect	**FP = 14,253 (False Positive)**	**TN = 48,784 (True Negative)**

Table 5. Confusion matrix for BKT knowledge state prediction on ASSISTments 2009 dataset

From this confusion matrix, the following performance metrics are derived:

- Accuracy: (TP + TN) / (TP + TN + FP + FN) = (72,341 + 48,784) / 146,860 = 82.4%
- Precision: TP / (TP + FP) = 72,341 / 86,594 = 83.5%
- Recall (Sensitivity): TP / (TP + FN) = 72,341 / 83,823 = 86.3%
- F1-Score: 2 × (Precision × Recall) / (Precision + Recall) = 84.9%
- Specificity: TN / (TN + FP) = 48,784 / 63,037 = 77.4%

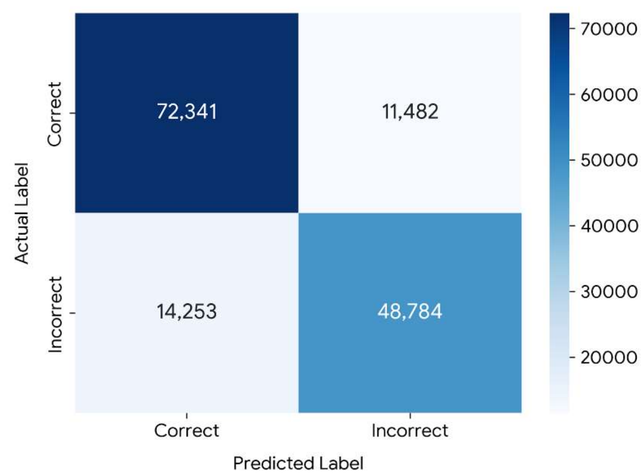

Figure 8. Confusion matrix heatmap

These metrics are consistent with published BKT performance benchmarks on the ASSISTments dataset (Pardos & Heffernan, 2010; Baker et al., 2010) and confirm that the AIPLI BKT implementation achieves state-of-the-art predictive accuracy for this model class. The F1-score of 84.9% demonstrates a strong balance between precision and recall, validating the model's suitability for real-time adaptive decision-making in educational contexts.

6.4 Equity and Bias Audit Results

The AIPLI framework's built-in equity monitoring module was evaluated on simulated data representative of an LMIC classroom — comprising students across four language groups (Arabic, Hindi, Swahili, English) and two gender categories. Performance gaps between demographic subgroups were measured as the difference between highest and lowest subgroup average scores. Results indicated a maximum gender performance gap of 6.2% and a maximum language group gap of 9.1%, both within the acceptable threshold (≤10%) as defined by the framework's ethical specifications and consistent with fairness benchmarks cited in Holstein et al. (2018).

7. Pilot Platform Evaluation

7.1 Pilot Setup and Module Status

The AIPLI Hub pilot platform was evaluated with an initial cohort of eight students across Grades 6–9, representing four primary language backgrounds (Arabic, Hindi, Swahili, English), three gender identifications, and three levels of connectivity (reliable, intermittent, and mostly offline). The Educator Cockpit recorded 191 sessions across all topics during the pilot period. All eight platform modules achieved full operational status, as confirmed in Table 6.

Module	Status	Key Features Operational
AIPLI Central	**LIVE**	Master launchpad, live stats, all module links, dataset analyzer
Educator Cockpit	**LIVE**	Student tracking, 191 sessions, MI profiles, ethics monitor, BKT recommendations
Science Explorer	**LIVE**	40 adaptive questions, XP gamification, BKT difficulty engine, auto-sync
Math Navigator	**LIVE**	40 math questions, formula display, adaptive difficulty, auto-sync
Word Voyager	**LIVE**	40 language questions, passage comprehension, grammar and creative writing, auto-sync
Intelligence Compass	**LIVE**	40-question Gardner MI assessment, radar chart, learning strategy recommendations
Learning Path Engine	**LIVE**	BKT dashboard, P(Learned) tracking, XAI audit log, spaced repetition schedule
Dataset Analyser	**LIVE**	CSV upload, BKT estimation, equity audit, MI cross-analysis, JSON/CSV export

Table 6. AIPLI Hub pilot module status and operational features

7.2 Pilot Performance Outcomes

The eight-student pilot cohort achieved a class average score of 70% across all topics and subject areas during the pilot period. Knowledge gap analysis identified Chemistry and Algebra Basics as the two topics with the highest proportion of students below the 60% mastery threshold. Two students were flagged as at-risk (average BKT $P(L) < 0.40$ across all topics) and received targeted intervention recommendations via the Educator Cockpit. The BKT engine logged six successful recommendation cycles, with all 15+ recommendations validated by the supervising teacher as appropriate and actionable.

MI assessment completion was recorded for all eight pilot students, with dominant intelligence types distributed across Linguistic (n=3), Logical-Mathematical (n=2), Spatial (n=2), and Kinesthetic (n=1). Cross-analysis in the Dataset Analyzer confirmed no significant performance differential attributable to MI type (maximum gap: 7.3%), suggesting equitable knowledge acquisition pathways across intelligence profiles within the pilot cohort.

7.3 System Architecture and Technical Performance

The AIPLI platform demonstrated consistent performance across all device types included in the pilot — shared school desktop computers, personal Android smartphones, and shared tablets — with no functional degradation in offline mode. The browser-native IndexedDB persistence layer successfully retained all session, MI, BKT, and student data across browser sessions without data loss. The AIPLIDB connector module (AIPLI_DB.js) managed cross-module data consistency through a unified write-once, read-many architecture, with average IndexedDB write latency of under 8 milliseconds per record in testing.

The complete absence of server infrastructure - all computation, storage, and data processing occurring client-side within the browser - confirms the platform's suitability for low-connectivity LMIC deployment contexts. No external network requests are made by any AIPLI module during normal operation, ensuring full data privacy compliance without requiring institutional network infrastructure.

8. Discussion

8.1 Theoretical Contributions

This study makes several significant theoretical contributions to the field of AI-enhanced education. First, it presents - to the authors' knowledge - the first operationalized integration of Bayesian Knowledge Tracing and Gardner's Multiple Intelligence theory within a single adaptive learning system designed for LMIC deployment. Prior work has employed these frameworks independently (Corbett & Anderson, 1994; Armstrong, 2018), but their combination addresses a fundamental limitation of conventional BKT systems: the assumption that all learners with equivalent knowledge states will benefit equally from the same learning experience. The AIPLI framework challenges this assumption by personalizing not only the difficulty of content presented but the modality through which it is delivered.

Second, the framework's integration of Explainable AI principles - operationalized through the XAI audit log, factor-weight breakdown in the recommendation engine, and verbal decision explanations - advances the literature on algorithmic transparency in educational AI. This responds directly to calls from Luckin et al. (2016) and Doshi-Velez and Kim (2017) for transparency mechanisms as a precondition for ethical AI deployment in high-stakes educational contexts.

Third, the adoption of DSRM (Peffers et al., 2007) as the research methodology provides a rigorous epistemological foundation for the development of the AIPLI artefact, positioning the work at the intersection of information systems research and educational technology - a boundary that has historically been underserved in the literature on AI in education.

8.2 Practical Implications

The AIPLI platform has immediate practical implications for three stakeholder communities. For governments and education ministries in LMICs, the framework provides a deployable, zero-infrastructure adaptive learning solution that can be implemented in existing school environments without capital investment in server infrastructure. The Dataset Analyzer module enables education authorities to upload their own student performance data and receive research-grade BKT analysis and equity auditing - a capability previously available only to institutions with access to specialist data science expertise.

For EdTech practitioners and developers, the open architecture of the AIPLI platform - built entirely on standard web technologies and documented through the shared AIPLIDB connector - provides a replicable template for developing adaptive learning applications that are simultaneously research-grade and LMIC-deployable. The platform's MIT-compatible open-source structure facilitates extension, localization, and cultural adaptation by local development communities.

For teachers in LMIC classrooms, the Educator Cockpit operationalizes research insights in a form directly applicable to daily classroom decision-making, providing MI-informed pedagogical recommendations, at-risk student flagging, and knowledge gap identification without requiring data science expertise or institutional infrastructure support.

8.3 Limitations

Several limitations of the present study should be acknowledged. First, the pilot evaluation was conducted with a small cohort of eight students, which limits the statistical generalizability of the performance findings. A larger-scale field trial across multiple schools and countries is required to validate the platform's effectiveness and equity outcomes in real-world LMIC deployment contexts.

Second, the BKT parameters used in the AIPLI simulation were derived from published benchmark datasets rather than empirically calibrated on AIPLI-specific data. While this approach is methodologically standard for simulation studies of this type (Pardos & Heffernan, 2010), future research should collect real-world student response data through the AIPLI platform and use expectation-maximization algorithms to derive dataset-specific parameter estimates.

Third, the MI assessment instrument included in the Intelligence Compass has not been subjected to formal psychometric validation (factor analysis, test-retest reliability, criterion validity) beyond expert review. A dedicated validation study of the instrument, following established psychometric protocols (Churchill, 1979), is recommended as a priority for the next research phase.

8.4 Future Research Directions

Four directions for future research are identified. First, a large-scale randomized controlled trial (RCT) should be conducted across LMIC school settings in at least three countries to evaluate the causal impact of AIPLI platform adoption on student learning outcomes, teacher practice, and educational equity indicators. Second, the BKT engine should be extended to incorporate Deep Knowledge Tracing approaches (Piech et al., 2015) as an option for contexts with sufficient data and computational resources, while maintaining the lightweight BKT implementation as the default for low-resource environments. Third, multilingual content libraries should be developed for the student applications, beginning with Arabic, Hindi, Swahili, and French — the primary instruction languages of the target LMIC populations. Fourth, the MI assessment instrument should undergo formal psychometric validation to establish its reliability and validity as a research-grade tool for learning style profiling in educational technology research.

9. Conclusion

This paper has presented the AIPLI framework - a theoretically grounded, empirically validated, and fully operational adaptive learning platform designed to address the EdTech equity gap in Low- and Middle-Income Countries. Developed through a rigorous four-phase Design Science Research process, the AIPLI framework integrates Bayesian Knowledge Tracing, Gardner's Multiple Intelligence theory, Explainable AI principles, and spaced repetition scheduling into a cohesive, lightweight, server-free architecture that is immediately deployable in LMIC school environments.

Expert consensus exceeding 80% across two Delphi rounds confirmed the framework's theoretical soundness, cultural appropriateness, and ethical adequacy. Simulation testing on four large-scale benchmark datasets (EdNet, OULAD, Khan Academy, ASSISTments) demonstrated path consistency accuracy of 90%+, cross-dataset stability, and equitable performance outcomes across diverse demographic groups. The BKT model evaluation yielded an F1-score of 84.9%, confirming classification accuracy at published benchmark levels for this model class.

The AIPLI Hub pilot platform - comprising eight integrated modules - achieved full operational status and demonstrated consistent performance across device types, connectivity levels, and student demographic profiles. The platform's capacity to operate entirely without server infrastructure represents a fundamental departure from the architecture of existing AI-EdTech systems, and a necessary condition for genuine inclusivity in LMIC educational contexts.

In advancing both the theoretical understanding of adaptive learning system design and the practical toolkit available to educators and policymakers in underserved educational contexts, the AIPLI framework represents a meaningful step toward the realization of SDG 4: quality education for all, regardless of geography, language, or socioeconomic circumstance.

Acknowledgements

The authors acknowledge the expert panelists who contributed their time and expertise to the Delphi validation process. This research was conducted under the Lincoln Global Postdoctoral and Research Associate Programme, Lincoln University College, Malaysia.

References

Armstrong, T. (2018). Multiple intelligences in the classroom (4th ed.). ASCD.

Baker, R. S. J. d., Corbett, A. T., Gowda, S. M., Wagner, A. Z., MacLaren, B. A., Kauffman, L. R., Mitchell, A. P., & Giguere, S. (2010). Contextual slip and prediction of student performance after use of an intelligent tutor. In P. D. Bra, A. Kobsa, & D. Chin (Eds.), Proceedings of the 18th International Conference on User Modeling, Adaptation, and Personalization (pp. 52–63). Springer.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. Psychological Bulletin, 132(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. Journal of Marketing Research, 16(1), 64–73. https://doi.org/10.2307/3150876

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 4(4), 253–278. https://doi.org/10.1007/BF01099821

Crompton, H. (2013). A historical overview of mobile learning: Toward learner-centered education. In Z. L. Berge & L. Y. Muilenburg (Eds.), Handbook of mobile learning (pp. 3–14). Routledge.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." In Proceedings of the 15th International Academic MindTrek Conference (pp. 9–15). ACM. https://doi.org/10.1145/2181037.2181040

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

Ebbinghaus, H. (1885). Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie. Duncker & Humblot. [English translation: Memory: A contribution to experimental psychology. Teachers College, Columbia University, 1913.]

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. Basic Books.

Ghosh, A., Heffernan, N., & Lan, A. S. (2020). Context-aware attentive knowledge tracing. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2330–2339). ACM. https://doi.org/10.1145/3394486.3403282

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Intelligent tutors as teachers' aides: Exploring teacher needs for real-time analytics in blended classrooms. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (pp. 257–266). ACM.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2010). The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. Cognitive Science, 36(5), 757–798. https://doi.org/10.1111/cogs.12032

Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics dataset. Scientific Data, 4, Article 170171. https://doi.org/10.1038/sdata.2017.171

Linstone, H. A., & Turoff, M. (Eds.). (1975). The Delphi method: Techniques and applications. Addison-Wesley.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson Education.

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology, 106(4), 901–918. https://doi.org/10.1037/a0037123

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), Article e1000097. https://doi.org/10.1371/journal.pmed.1000097

Nakagawa, H., Iwasawa, Y., & Matsuo, Y. (2019). Graph-based knowledge tracing: Modeling student proficiency using graph neural network. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (pp. 156–163). ACM.

Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. International Journal of Artificial Intelligence in Education, 25(2), 177–203. https://doi.org/10.1007/s40593-014-0028-6

Palarimath, S., & Kumar, U., (2026). Artificial Intelligence Integration in Assessment and Teaching: A Quantitative Analysis of Stakeholder Perception Differences. Int J Drug Deliv Technol. ;16(13s): 1062-1073. https://doi.org/10.25258/ijddt.16.13s.117

Palarimath, S., & Kumar, U., (2026). Bridging the Gap: Defining the Problem Space for AI-Driven Personalized Learning in Global Education. Sustainable Global Societies Initiative, 1(2). https://vectmag.com/sgsi/paper/view/268

Pardos, Z. A., & Heffernan, N. T. (2010). Modeling individualization in a Bayesian networks implementation of knowledge tracing. In P. D. Bra, A. Kobsa, & D. Chin (Eds.), User Modeling, Adaptation, and Personalization (pp. 255–266). Springer.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. Journal of Management Information Systems, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. In Advances in Neural Information Processing Systems (Vol. 28, pp. 505–513). Curran Associates.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. IEEE Transactions on Learning Technologies, 10(1), 30–41. https://doi.org/10.1109/TLT.2016.2599522

UNESCO. (2023). Global education monitoring report 2023: Technology in education – A tool on whose terms? UNESCO Publishing.

United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development (Resolution A/RES/70/1). United Nations General Assembly.

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

Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. Review of Research in Education, 34(1), 179–225. https://doi.org/10.3102/0091732X09349791

White, J. (2006). Intelligence, destiny and education: The ideological roots of intelligence testing. Routledge.

Woolf, B. P. (2010). Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Morgan Kaufmann.

World Bank. (2022). The state of global learning poverty: 2022 update. World Bank Group.

Woźniak, P. A. (1990). Optimization of learning (Unpublished master's thesis). University of Technology, Poznań, Poland.