REVISTA de PEDAGOGIE

JOURNAL of PEDAGOGY

http://revped.ise.ro

Print ISSN 0034-8678; Online ISSN: 2559 - 639X

THE USE OF AI TO OPTIMISE BACCALAUREATE OUTCOMES AN EDUCATIONAL DECISION-MAKING: TOWARDS AN INTEGRATIVE PEDAGOGICAL FRAMEWORK FOR HIGH-STAKES EXAMS

Utilizarea IA pentru optimizarea rezultatelor bacalaureatului şi a procesului decizional în domeniul educaţiei: către un cadru pedagogic integrativ pentru examenele cu miză mare

Olimpius ISTRATE, Bogdan CRISTESCU

Journal of Pedagogy, 2025 (2), 225 - 249

https://doi.org/10.26755/RevPed/2025.2/225

The online version of this article can be found at: https://revped.ise.ro/category/2025/

Published by:

INSTITUTUL DE ŞTIINŢE ALE EDUCAŢIEI

http://www.ise.ro/

Further information about *Revista de Pedagogie – Journal of Pedagogy* can be found at:
Editorial Policy: http://revped.ise.ro/editorial-policy/
Author Guidelines: http://revped.ise.ro/en/author-guidelines/

Revista de Pedagogie - Journal of Pedagogy, 2025 (2), 225 – 249
https://doi.org/10.26755/RevPed/2025.2/225

THE USE OF AI TO OPTIMISE BACCALAUREATE OUTCOMES AND EDUCATIONAL DECISION-MAKING. TOWARDS AN INTEGRATIVE PEDAGOGICAL FRAMEWORK FOR HIGH-STAKES EXAMS

Olimpius Istrate*

University of Bucharest
Faculty of Psychology and Education Sciences
Office of the Secretary-General of the European Schools
European Baccalaureate Unit
Brussels, Belgium
olimpius.istrate@eursc.eu

Bogdan Cristescu**

National Centre for Curriculum and Evaluation in Education
Bucharest, Romania
bogdan.cristescu@rocnee.ro

Abstract
High-stakes examinations like the baccalaureate represent critical junctures in students' academic and life trajectories, yet current assessment preparation and evaluation often fail to provide actionable insights for educational improvement. Today, digital learning and assessment platforms are able to capture rich behavioural data such as response times and answer revision patterns, providing insights beyond

* Senior lecturer, PhD, University of Bucharest, Faculty of Psychology and Education Sciences, Bucharest, Romania. Office of the Secretary-General of the European Schools, European Baccalaureate Unit, Brussels, Belgium.
** Head of the Curriculum, Evaluations, and National Assessments Department, National Centre for Curriculum and Evaluation in Education, Bucharest, Romania.

traditional analyses. However, the existing practices typically focus on isolated aspects of the examination process, not yet leveraging AI comprehensively across preparation, implementation, and post-exam analysis phases.
This paper introduces the APIA framework (Assessment Preparation, Implementation, and Analysis), which takes into account the possibilities offered by AI. The framework connects three critical phases through systematic AI integration. It operates across three dimensions: processual (distinguishing preparation, implementation, and analysis), stakeholder (mapping roles from students to policymakers), and cognitive-taxonomic (aligning AI capabilities with Bloom's taxonomy levels). Phase one leverages AI for personalized student preparation through adaptive learning systems. Phase two employs predictive modelling using longitudinal data to enable proactive interventions at individual, institutional, and systemic levels. Phase three transforms examination results into actionable intelligence for curriculum development, teacher training, and policy improvement. AI-enhanced APIA model prioritizes pedagogical coherence, addressing implementation challenges through explicit stakeholder role definition and cognitive-taxonomic alignment that prevents AI clustering at basic or advanced levels, while neglecting intermediate thinking processes. This represents an initial attempt to bridge the gap between isolated AI applications and comprehensive educational improvement, offering practitioners a structured approach to systematic AI integration in assessment contexts.

Keywords: artificial intelligence in education, baccalaureate, digital pedagogy, evidence-based education, high-stakes exams.

Rezumat
Examenele cu miză mare precum bacalaureatul reprezintă momente critice în parcursurile academice, personale şi profesionale ale elevilor, însă, în prezent, pregătirea şi evaluarea examinării eşuează adesea în a furniza sugestii de ameliorare. Platformele digitale de învăţare şi evaluare sunt capabile astăzi să surprindă date complexe despre comportamentul în examen, precum timpii de răspuns şi tiparele de revizuire a răspunsurilor, oferind perspective care depăşesc analizele tradiţionale. Cu toate acestea, practicile existente se concentrează de obicei asupra aspectelor izolate ale procesului de examinare, fără a valorifica încă integrarea posibilităţilor oferite de inteligenţa artificială (IA) în fazele de pregătire, implementare şi analiză post-examen.
Această lucrare propune un cadru pentru etapele evaluării care ia în considerare posibilităţile IA. Acesta conectează trei faze critice – APIA (Assessment Preparation, Implementation, and Analysis), operând pe trei dimensiuni: procesual (distinge pregătirea, implementarea şi analiza), părţile interesate (cartografiază

rolurile, de la elevi la factorii de decizie) şi cognitiv-taxonomic (aliniază posibilităţile IA cu nivelurile taxonomiei lui Bloom). Prima fază valorifică IA pentru pregătirea personalizată a elevilor prin sisteme de învăţare adaptivă. A doua fază utilizează modelarea predictivă folosind date longitudinale pentru a permite intervenţii proactive la nivel individual, instituţional şi sistemic. A treia fază transformă rezultatele examenelor în informaţie utilă pentru dezvoltarea curriculumului, pentru formarea cadrelor didactice şi pentru ameliorarea politicilor educaţionale.
Modelul APIA îmbunătăţit prin apelul la IA prioritizează coerenţa pedagogică, abordând provocările de implementare prin definirea explicită a rolurilor părţilor interesate şi alinierea cognitiv-taxonomică care previne gruparea IA doar la nivelul de bază sau avansat, neglijând capacităţile cognitive de nivel intermediar. Acesta reprezintă o primă încercare de a construi o punte între diverse aplicaţii izolate ale IA şi ameliorarea educaţională care necesită o abordare holistică, oferind practicienilor o perspectivă structurată pentru integrarea sistematică a IA în contexte de evaluare.

***Cuvinte-cheie**: bacalaureat, educaţie bazată pe dovezi, examene cu miză mare, inteligenţa artificială în educaţie, pedagogie digitală.*

1. Introduction

Despite of the importance of high-stakes examinations for students' academic pathways, current approaches to preparation, administration, and post-examination analysis often operate in isolation, failing to harness the full potential of available data and technological capabilities to optimize both individual outcomes and systemic educational quality. This fragmentation has resulted in missed opportunities for proactive intervention, personalized support, and systematic improvement that could fundamentally transform how educational systems approach assessment and learning optimization.

The emergence of artificial intelligence technologies presents unprecedented opportunities to address these longstanding challenges through comprehensive integration across the entire examination ecosystem. Recent advances in machine learning, natural language processing, and educational data mining have demonstrated remarkable capabilities in predicting student performance, providing adaptive learning support, and generating insights from complex

educational datasets. However, these technological innovations have remained largely compartmentalized, with most implementations focusing on isolated aspects of the assessment process rather than creating systematic connections between preparation and post-examination analysis phases. This isolation limits the transformative potential of AI in education, often reducing sophisticated technologies to simple automation tools rather than leveraging their capacity for comprehensive educational enhancement.

This paper proposes a paradigm shift toward integrated AI implementation in high-stakes examinations through the development of a pedagogical framework that synthesizes artificial intelligence applications across three critical phases: preparation, implementation, and analysis. The AI-enhanced APIA framework (Assessment Preparation, Implementation, and Analysis) represents an initial systematic attempt to create coherent connections between AI-powered student preparation systems, predictive modelling for proactive interventions, and post-examination analysis for systemic improvement. By addressing the fragmentation that has characterized previous approaches, this framework seeks to transform high-stakes examinations from isolated assessment events into dynamic components of a continuous learning optimization cycle that serves both individual student success and broader educational improvement objectives.

2. Approaches and Sources of Data for Evidence-Based Decision-Making at Education System Level

The transformation of educational decision-making from intuition-based practices toward evidence-informed approaches is a significant shift in how educational systems operate and improve. This evolution has been driven by the proliferation of digital technologies and the increasing availability of diverse educational data sources, creating opportunities for more systematic and effective decision-making processes at all levels of the education system.

2.1. The Nature of Educational Data

Educational data encompasses “any information systematically collected, recorded, and used to inform teaching, learning, and decision-making

processes" (Brazauskienė, 2025). This broad definition includes both quantitative and qualitative forms, ranging from numerical records to descriptive insights that capture the complexity of educational environments. Contemporary educational data can be categorized into four primary types (Hamilton et al., 2009; Mandinach, 2012):

- *Student-level data* includes demographic information, academic performance records, learning behaviours and engagement metrics, and social-emotional indicators.
- *Instructional data* covers curriculum content, teaching methods, assessment tasks, and lesson observations.
- *Institutional data* encompasses school resources, staffing records, and enrolment statistics.
- *Contextual data* captures family backgrounds, policy frameworks, and societal factors influencing learning.

The digital era has significantly expanded data collection capabilities through learning management systems and e-learning platforms, which capture traditional academic metrics alongside behavioural patterns and engagement data (Gašević et al., 2017; Kostopoulos & Kotsiantis, 2022; Thompson & Sellar, 2018; Torrisi-Steele et al., 2019). Contemporary systems increasingly incorporate multimodal inputs including audio, video, and biometric data, enabling more comprehensive analysis of learning behaviours and emotional states during educational activities. This technological expansion enables real-time monitoring and more nuanced understanding of learning processes, though it simultaneously raises critical questions about data privacy, student consent, and the ethical implications of pervasive monitoring in educational environments.

2.2. Three Approaches to Data-Based Decision Making

The evolution of evidence-based decision-making in education has produced three distinct but interconnected approaches that differ primarily in how they balance data analysis with professional judgment and contextual considerations (Brazauskienė, 2025). Understanding these distinctions is important for educational leaders seeking to implement effective data use practices that align with their institutional contexts and educational goals.

Data-driven decision making (DDDM) represents the most deterministic approach, where decisions are made primarily on the basis of objective, systematically gathered data rather than intuition or personal experience (Brazauskienė, 2025). This approach assumes that data provides an objective foundation for decision-making that minimizes subjective bias and enhances accountability (Mandinach, 2012).

The DDDM process follows a systematic cycle: data collection, analysis, decision implementation, and monitoring. Technology plays a central role, with learning analytics, prediction tools, and real-time data systems enabling sophisticated analysis of student performance and engagement patterns. However, this approach faces criticism for potentially overshadowing professional expertise and contextual knowledge, particularly when complex educational decisions require nuanced understanding beyond what data alone can provide.

Data-based decision making (DBDM) represents a more collaborative and systematic approach that, while grounding decisions in data, incorporates broader processes of problem identification, goal setting, and stakeholder collaboration. Unlike DDDM, DBDM explicitly acknowledges the need for professional interpretation and contextual factors in decision-making processes (Van der Kleij et al., 2015).

The DBDM cycle encompasses problem definition, data source selection, collaborative analysis among educators and administrators, and implementation with ongoing monitoring. This approach has shown particular effectiveness in supporting students with special needs, where data must be combined with advanced specialised pedagogical expertise to develop appropriate interventions. However, successful implementation requires substantial investment in data literacy training for educators and administrative support systems (Cramer et al., 2014; Decabooter et al., 2024).

Data-informed decision making (DIDM) is the most balanced approach, positioning data as one valuable input alongside professional judgment, pedagogical experience, and local knowledge. This approach recognizes that data provides necessary information but achieves its real value only when synthesized with educators' professional expertise and contextual understanding (Wang, 2021).

DIDM integrates quantitative data with qualitative insights, acknowledging that educational contexts are complex and that effective decisions require understanding of local cultures, student needs, and community dynamics. The approach explicitly recognizes that academic achievement data alone provides an incomplete picture of student learning, as multiple factors including stress levels, emotional states, and social circumstances significantly influence educational outcomes. Emerging technologies, including machine learning algorithms and generative artificial intelligence, can enhance this process by analysing the complex interactions between these contextual factors and calculating potential impacts of different decision scenarios. Advanced AI-powered systems can incorporate multimodal data from sensors monitoring physiological indicators, learning behaviours, and social interactions to provide comprehensive support that addresses not only academic needs but also social and emotional well-being. However, these technological capabilities serve to inform rather than replace educator agency, providing sophisticated analytical support for human decision-making (Lee & Lee, 2025).

The evolution toward data-informed decision-making in education creates both opportunities and challenges for implementing advanced technologies in assessment contexts. While DIDM principles provide the conceptual foundation for integrating professional judgment with empirical evidence, the practical application of these principles increasingly depends on sophisticated technological tools capable of processing complex educational data at scale. Artificial intelligence emerges as a particularly promising technology for enhancing data-informed decision-making processes, offering capabilities that surpass traditional analytics.

3. AI Applications in High-Stakes Educational Assessment

The integration of artificial intelligence into high-stakes educational assessment represents a paradigm shift from traditional evaluation methods toward sophisticated, data-driven systems capable of analysing complex learning behaviours and predicting educational outcomes with unprecedented accuracy. The research reviewed in this chapter spans multiple educational levels, from primary through higher education, reflecting the current state of AI implementation in assessment contexts. While high-stakes examination research at the secondary

level (including baccalaureate) remains limited, the pedagogical principles and technological capabilities demonstrated across educational contexts provide valuable foundations for developing comprehensive frameworks applicable to various assessment scenarios. Where research specifically addresses secondary education or high-stakes examinations, this is explicitly noted.

3.1. Performance Prediction and Early Warning Systems

Contemporary AI applications in educational assessment have achieved remarkable success in predicting student performance, with machine learning models consistently demonstrating accuracy rates exceeding 90% when incorporating diverse features including academic history, demographic characteristics, behavioural patterns, and socio-economic factors (Akçapınar et al., 2019). Today's predictive capabilities extend beyond simple grade forecasting to encompass comprehensive early warning systems that identify students at risk of academic failure or dropout.

According to Akçapınar et al. (2019), the Classification and Regression Tree (CART) algorithm, supplemented by AdaBoost (adaptive boosting), has demonstrated over 90% accuracy across multiple datasets in predicting students' pass/fail status using time-dependent features derived from learning management systems. Similarly, Costa et al. (2019) show that advanced machine learning techniques have proven effective in reducing failure rates in challenging subjects, with Support Vector Machine (SVM) algorithms achieving 83% accuracy in early identification of students with high probability of failure in introductory programming courses.

The scope of predictive modelling extends across multiple educational levels, with research demonstrating that pedagogically grounded predictors, such as assessment-related data, provide more actionable insights than non-pedagogical predictors like socioeconomic status (Bulut et al., 2024). Student performance prediction systems have successfully utilized various machine learning approaches including Random Forest, Neural Networks, and k-Nearest Neighbor algorithms, achieving classification accuracies between 70-75% for final exam grade predictions using only midterm exam grades, department data, and faculty information (Yağcı, 2022).

3.2. Behavioural Analytics and Learning Trace Data

Digital learning and assessment platforms capture rich behavioural data that extends far beyond traditional academic metrics, providing insights into response times, answer revision patterns, and engagement behaviours that reveal deeper understanding of learning processes (Bulut et al., 2024). These multimodal data sources enable comprehensive analysis of student learning behaviours, incorporating audio, video, and biometric data to create holistic portraits of student competencies (Caspari-Sadeghi, 2022).

Learning analytics systems leverage this behavioural data to identify patterns in system access, navigational actions, and interaction behaviours that correlate with academic success or failure (Yağcı, 2022). The analysis of e-assessment logs has revealed significant connections between student behaviour and performance, enabling educators to move beyond traditional item analysis toward understanding the complex interplay between behavioural patterns and learning outcomes (Lahza et al., 2023).

3.3. Automated Assessment and Scoring Systems

AI-powered automated scoring systems have revolutionized the evaluation of complex assessments, particularly in essay scoring where natural language processing techniques achieve remarkable accuracy. Advanced automated essay scoring systems utilizing latent semantic analysis and machine learning algorithms have demonstrated accuracy rates of 97.6%, significantly outperforming reference models and approaching human-level consistency (Anjum, 2023).

The implementation of automated writing evaluation systems addresses traditional challenges of subjective scoring while providing consistent, scalable assessment capabilities. These systems employ sophisticated natural language processing to evaluate not merely keyword matching but semantic meaning and discourse coherence, enabling fair and reliable scoring of open-ended responses (Richardson & Clesham, 2021).

3.4. Multi-Level Predictive Modelling

Contemporary AI applications enable predictive modelling at multiple organizational levels, from individual students to entire educational cohorts. Educational Data Mining techniques can forecast various educational outcomes including enrolment patterns, dropout rates, and academic performance trajectories, providing actionable intelligence for proactive intervention strategies (Caspari-Sadeghi, 2022).

The predictive utility of assessment results extends beyond immediate academic outcomes to encompass broader educational pathways. Student Grade Point Averages in ninth grade can predict high school dropout likelihood, while formative assessment scores and task completion times serve as strong predictors of summative assessment performance (Bulut et al., 2024). These multi-level predictions enable targeted resource allocation and intervention strategies at individual, classroom, school, and system levels.

3.5. Current Limitations and Gaps

Current AI applications in high-stakes assessment, while demonstrating remarkable capabilities in specific domains, reveal critical implementation gaps when examined systematically. Predictive models excel at forecasting performance using complete datasets but falter with incomplete student records – a common reality in educational systems with high mobility rates. Automated scoring systems achieve near-human reliability for essay evaluation in controlled conditions (Anjum, 2023), yet their adoption in actual high-stakes examinations remains minimal due to stakeholder concerns about fairness and transparency. Behavioral analytics effectively identify at-risk patterns in digital learning environments (Yağcı, 2022), but these insights rarely inform proactive interventions before examination administration. These examples illustrate a broader pattern: most existing approaches focus on isolated aspects of the assessment process, addressing individual components rather than integrating AI capabilities across the complete examination lifecycle from preparation through post-assessment analysis (Fartuşnic et al., 2025).

The explainability challenge represents a critical limitation, as complex predictive models often function as "black boxes" that provide accurate predictions without transparent reasoning processes. The application of Explainable AI (XAI) becomes essential to establish trust between educators and AI systems, enabling human validation of prediction results before implementing interventions (Bulut et al., 2024).

Current systems often struggle with cultural and contextual factors that influence assessment validity. Regional differences in educational practices, data protection regulations, and cultural views on assessment create challenges for generalizing AI applications across diverse educational contexts (Hartmann, 2023). Furthermore, the integration of AI into high-stakes examinations must address fairness concerns, algorithmic bias, and the potential for circular reproduction of errors inherent in training data.

3.6. Toward Comprehensive Framework Integration

A fundamental limitation in current endeavour is that research specifically examining AI applications in secondary education high-stakes examinations, including baccalaureate-level assessment, remains scarce. This gap reflects both the higher risks associated with implementing experimental technologies in high-stakes contexts and the slower adoption of comprehensive digital assessment platforms at secondary levels. However, the technological capabilities and pedagogical principles demonstrated in higher education research provide important foundations for developing frameworks applicable to secondary education, even as these require adaptation to the distinct developmental, curricular, and stakeholder contexts of baccalaureate examinations.

The existing research demonstrates that while individual AI applications show remarkable success in specific domains, the field lacks specific frameworks that integrate these capabilities across the entire high-stakes examination ecosystem. The need for systematic approaches that combine predictive modelling, adaptive assessment, behavioural analytics, and post-examination analysis becomes increasingly apparent as educational systems seek to maximize both individual student outcomes and systemic improvements.

This gap in comprehensive integration provides the foundation for developing pedagogical frameworks that leverage AI capabilities across preparation, prediction, and analysis phases of high-stakes examinations. Such frameworks must address technological capabilities along with ethical considerations, stakeholder needs, and (mostly) pedagogical coherence to ensure that AI serves educational goals rather than driving them.

4. A Pedagogical Framework for AI-Enhanced Assessment Preparation, Implementation, and Analysis (APIA)

To address the relationship between technological capabilities and systematic implementation, we propose a multi-dimensional framework, which synthesizes AI applications across all examination phases.

4.1. From Classical to AI-Enhanced Assessment Paradigms

In the AI era, assessment practices in education are undergoing significant evolution, with emerging applications in both formative and summative evaluation contexts. This evolution reflects ongoing efforts to move beyond traditional assessment-for-accountability models – where examinations primarily serve to measure and report student achievement for external stakeholders – toward more process-oriented approaches that leverage technological capabilities while preserving essential human expertise. The digital systems offer potential for more continuous and adaptive assessment experiences. However, the extent to which AI-enhanced systems can deliver on promises of truly personalized, continuous assessment remains contested, with implementation largely confined to small-scale studies and pilot programs rather than systematic deployment in high-stakes examination contexts. While some researchers anticipate transformative potential (Williams, 2023), others caution that AI applications have not yet demonstrated sufficient legitimacy and pertinence for widespread employment in high-stakes examinations, pointing to unresolved issues of validity, fairness, and pedagogical coherence (Bulut et al., 2024; Fartuşnic et al., 2025).

The emerging AI-enhanced paradigm fundamentally reconceptualizes assessment as a continuous, integrated component of the learning process rather than an isolated evaluative event. Formative assessment within this new paradigm becomes process-oriented, dialectic, and confidence-building, supporting scaffolded use of AI technologies that enhance rather than replace human judgment (Fartuşnic et al., 2025). This approach enables real-time monitoring of learning progress, immediate feedback provision, and adaptive responses to individual student needs, creating what Williams (2023) describes as Assessment for Learning that is "highly compatible with the employment of trace data" and capable of "providing real-time formative feedback".

Summative assessment in the AI era shows potential for evolving toward competency-oriented evaluation that integrates knowledge, skills, and attitudinal dimensions – a shift already underway in baccalaureate examinations for over three decades, accelerated by digital technologies and now potentially enhanced by AI capabilities. This ongoing transformation requires that high-stakes examinations continue adapting their assessment design to capture the full spectrum of competencies rather than isolated skills or knowledge recall. Specifically, AI integration could support this evolution through several concrete mechanisms: enabling the analysis of complex problem-solving processes across multiple assessment tasks to infer higher-order competencies; facilitating the evaluation of authentic performance that demonstrates knowledge application in realistic contexts; and identifying patterns in student responses that reveal not only what students know but how they approach unfamiliar problems and adapt their strategies. However, such applications require careful design to ensure that AI tools support rather than constrain the assessment of comprehensive competencies, avoiding the risk of narrowing evaluation to easily automated components at the expense of the holistic, multidimensional nature of competency-based assessment that contemporary baccalaureate examinations aim to achieve.

4.2. The Missing Link: Contextual Interpretation

A critical limitation of current educational data systems lies in their insufficient capacity for contextual interpretation of assessment results. While traditional approaches excel at collecting quantitative performance metrics, they often

fail to provide the nuanced understanding necessary for effective educational decision-making. As Brazauskienė (2025) emphasizes in her analysis of data-informed decision-making, the real value of educational data emerges only when synthesized with educators' professional expertise and contextual understanding.

AI technologies offer unprecedented potential for filling this missing link by enabling sophisticated analysis of the complex interactions between multiple contextual factors and calculating potential impacts of different decision scenarios (Lee & Lee, 2025). Advanced AI-powered systems can incorporate multimodal data from various sources monitoring physiological indicators, learning behaviours, and social interactions to provide comprehensive support that addresses academic, social, and emotional dimensions of student development. However, these technological capabilities serve to inform rather than replace educator agency, providing sophisticated analytical support for human decision-making processes.

4.3. The AI-enhanced APIA framework – Assessment Preparation, Implementation, and Analysis

The AI-enhanced APIA framework represents an integrative approach that attempts to synthesize AI applications across three critical phases of the high-stakes examination, each building upon the data-informed decision-making principles established in Chapter 1. This model advances beyond the isolated applications identified in Chapter 2 by creating systematic connections between preparation, prediction, and post-examination analysis phases, ensuring that AI serves pedagogical objectives rather than constraining educational practices.

The framework incorporates three essential dimensions that work synergistically to optimize baccalaureate outcomes:

(1) a processual dimension that distinguishes between preparation, implementation, and utilization phases;
(2) a stakeholder dimension that maps roles and responsibilities across the educational ecosystem; and
(3) a cognitive-taxonomic dimension that aligns AI capabilities with different levels of learning complexity.

This multidimensional approach ensures that technological innovations serve educational objectives, preserving human expertise in areas requiring judgment, creativity, and ethical reasoning (Fartuşnic et al., 2025).

***Table no. 1**. A three-dimensional architecture for the AI-enhanced APIA framework – Assessment Preparation, Implementation, and Analysis*

Dimension	Focus	Key Components	Primary Function
Processual	*What* - Assessment phases and workflows	Preparation → Implementation → Utilization	Distinguishes between different phases of assessment lifecycle
Stakeholder	*Who* - Roles and responsibilities	Students, Teachers, Administrators, Policymakers, Researchers	Maps accountability and agency across educational ecosystem
Cognitive-taxonomic	*How* - Cognitive complexity levels	Bloom's Taxonomy integration (Remember → Create)	Aligns AI capabilities with learning complexity

Phase 1: AI-Powered Student Preparation Systems

The first phase of the AI-enhanced APIA focuses on leveraging AI technologies to enhance individual student preparation through adaptive, personalized learning experiences that respond dynamically to knowledge gaps and learning patterns. This phase transforms traditional "one-size-fits-all" preparation approaches by implementing sophisticated algorithms that generate baccalaureate-style practice assessments tailored to each student's specific needs and learning pathway.

AI-powered preparation systems utilize comprehensive analysis of student performance data, behavioural patterns, and engagement metrics to create individualized learning pathways that optimize preparation efficiency. These systems can identify conceptual misunderstandings through pattern analysis across multiple assessment attempts, detect correlations between performance on different question types, and recognize learning progression trends that inform adaptive content delivery (Owan et al., 2023). The continuous feedback loops embedded within these systems ensure that preparation activities remain aligned with individual learning needs while maintaining connection to baccalaureate standards and expectations.

***Table no. 2**. AI-enhanced APIA, phase one: AI-Powered Student Preparation Systems*

Component	AI Application	Key Features	Expected Outcomes
Adaptive learning pathways	Personalized practice test generation	• Dynamic difficulty adjustment • Knowledge gap identification • Learning pattern analysis	• Optimized preparation efficiency • Individualized learning trajectories
Behavioural analytics	Response pattern analysis	• Time-on-task monitoring • Engagement tracking • Performance correlation analysis	• Early intervention identification • Learning strategy optimization
Real-time feedback	Immediate performance guidance	• Instant error correction • Conceptual scaffolding • Progress visualization	• Continuous learning improvement • Enhanced self-efficacy

Phase 2: Predictive Modelling for Proactive Interventions

The second phase harnesses longitudinal academic data from lower and upper secondary education to forecast baccalaureate performance at multiple organizational levels – individual students, classes, schools, and educational cohorts – enabling proactive interventions and strategic resource allocation. This predictive capability extends beyond simple performance forecasting to encompass comprehensive early warning systems that identify students at risk of academic difficulties while there remains time for effective intervention.

Machine learning algorithms within this phase process diverse data sources including academic history, behavioural patterns, engagement metrics, and contextual factors to generate accurate performance predictions that inform decision-making at multiple levels (Bulut et al., 2024). The multi-level nature of these predictions enables targeted interventions that address individual student needs while simultaneously informing institutional resource allocation and policy development decisions. Explainable AI principles ensure that predictive insights remain transparent and actionable, enabling human validation of algorithmic recommendations before implementation.

***Table no. 3**. AI-enhanced APIA, phase two: Predictive Modelling for Proactive Interventions*

Component	AI Application	Key Features	Expected Outcomes
Multi-level prediction	Longitudinal data analysis	• Individual student forecasting • Class-level predictions • School/ system analysis	• Proactive intervention targeting • Strategic resource allocation
Early warning systems	Risk identification algorithms	• Academic failure prediction • Engagement decline detection • Performance trajectory analysis	• Timely support provision • Prevention-focused approach
Intervention optimization	Evidence-based recommendation	• Personalized support strategies • Resource allocation guidance • Success probability modelling	• Improved intervention effectiveness • Data-driven decision making

Phase 3: Post-Examination Analysis for Systemic Improvement

The final phase transforms baccalaureate examination results into actionable intelligence for systemic educational improvement through sophisticated analysis of regional, institutional, and demographic performance patterns. This phase moves beyond traditional result reporting to provide comprehensive insights that inform targeted improvements in teacher training, curriculum development, and student support services.

AI-powered post-examination analysis systems can identify specific performance patterns across different demographic groups, geographic regions, and institutional contexts, revealing systematic strengths and challenges that might remain hidden in traditional analysis approaches. These insights enable evidence-based improvements in educational policies and practices, creating feedback loops that continuously enhance the effectiveness of the entire educational system. The integration of multiple data sources and analytical perspectives ensures that systemic improvements address root causes rather than merely symptoms of educational challenges.

***Table no. 4**. AI-enhanced APIA, phase three: Post-Examination Analysis for Systemic Improvement*

Component	AI Application	Key Features	Expected Outcomes
Performance pattern analysis	Multi-dimensional data mining	• Regional comparison analysis • Demographic equity assessment • Institutional effectiveness evaluation	• Targeted improvement strategies • Evidence-based policy development
Curriculum optimization	Content effectiveness analysis	• Learning objective alignment • Knowledge gap identification • Instructional sequence optimization	• Enhanced curriculum design • Improved learning outcomes
Professional development	Teacher training needs analysis	• Instructional effectiveness patterns • Professional development targeting • Best practice identification	• Enhanced teaching quality • Systematic capacity building

4.4. A Stakeholder Integration Matrix

The Stakeholder Integration Matrix represents another important implementation component of the APIA framework that addresses a fundamental challenge identified in educational research: the frequent disconnect between technological capabilities and successful institutional adoption. This matrix emerged from the analysis of technology implementation studies that revealed how misaligned roles and unclear responsibilities across the educational ecosystem consistently undermined computer-assisted assessment initiatives, even when the underlying technology demonstrated strong performance. By explicitly mapping the distinct responsibilities, engagement levels, and key concerns of students, teachers, administrators, policymakers, and researchers, the matrix provides practitioners with a structured roadmap for collaborative implementation that prevents the technological determinism often observed in failed AI adoption attempts.

***Table no. 5**. AI-enhanced APIA: Stakeholder Integration Matrix*

Stakeholder	Primary Roles	AI Interaction Level	Key Responsibilities
Students	Active learners, assessment takers	*High* - Direct interaction with AI tools	• Engage with adaptive systems • Provide learning data • Utilize AI feedback
Teachers	Instructional design, assessment implementation	*High* - AI-assisted instruction and evaluation	• Interpret AI insights • Guide student learning • Validate AI recommendations
Administrators	Resource allocation, policy implementation	*Medium* - AI-informed decision making	• Allocate AI tools and training • Monitor implementation effectiveness • Ensure ethical AI use
Policymakers	Framework development, systemic oversight	*Medium* - AI-informed policy development	• Establish AI governance frameworks • Ensure equity and fairness • Support infrastructure development
Researchers	Framework validation, continuous improvement	*High* - AI system development and evaluation	• Develop AI algorithms • Validate prediction models • Ensure scientific rigor

The matrix serves as a diagnostic tool for identifying stakeholder alignment gaps. It is also useful as a planning instrument for designing implementation strategies that balance innovation with pedagogical integrity, ensuring that each stakeholder group can contribute their unique expertise, maintaining accountability for their respective domains within the comprehensive AI-enhanced educational ecosystem.

4.5. Cognitive-Taxonomic Alignment

The Cognitive-Taxonomic Alignment part of the model addresses a fundamental tension identified in current research on AI applications in educational assessment: the tendency for artificial intelligence tools to cluster at either basic cognitive levels or advanced creative tasks, while neglecting the complex middle ranges of analytical and evaluative thinking that characterize much of meaningful learning. This clustering phenomenon directly impacts assessment construction, as test developers often gravitate toward either straightforward knowledge-recall items (easily automated through AI) or open-ended creative tasks (where AI provides generative

support), inadvertently creating gaps in the systematic assessment of intermediate cognitive processes such as analysis, synthesis, and evaluation.

The alignment framework emerges from systematic analysis revealing that contemporary AI assessment applications often fail to capture the full spectrum of cognitive processes embedded in modern curricula, potentially narrowing educational focus to easily automated tasks rather than supporting comprehensive intellectual development (Du et al., 2025; Jose et al., 2025; Lubbe et al., 2025; Stokkink, 2025). From an item construction perspective, this limitation constrains educators' ability to design assessments that comprehensively measure learning objectives across all taxonomic levels, forcing artificial choices between AI-supported efficiency and cognitive complexity. The framework demonstrates how AI can provide cognitive scaffolding across Bloom's hierarchical taxonomy, enabling test developers to design items at appropriate cognitive levels, aligned with curricular objectives and educational standards.

***Table no. 6**. AI-enhanced APIA: A Proposal for Cognitive-Taxonomic Alignment*

Bloom's Level	Human Capabilities	AI Support Functions	Possible Integration Approach
Remember	Contextual recall, meaningful connections	Pattern recognition, fact retrieval	AI handles routine recall, humans provide context
Understand	Conceptual interpretation, meaning-making	Content analysis, explanation generation	AI identifies patterns, humans interpret significance
Apply	Situated problem-solving, transfer	Solution recommendation, procedure guidance	AI suggests approaches, humans adapt to context
Analyse	Critical evaluation, relationship identification	Data mining, correlation analysis	AI processes data, humans provide critical judgment
Evaluate	Judgment formation, criteria application	Evidence synthesis, bias detection	AI provides comprehensive analysis, humans make judgments
Create	Innovation, original synthesis	Idea generation support, creative scaffolding	AI offers creative support, humans drive innovation

In practice, this alignment serves as both a design tool for educators developing AI-enhanced assessments and a quality assurance mechanism for ensuring that technological integration maintains pedagogical integrity by supporting rather than supplanting the full range of cognitive skills essential for deep learning and meaningful intellectual development.

4.6. Ethical and Implementation Considerations

The comprehensive implementation of the APIA framework requires careful attention to ethical considerations, including data privacy, algorithmic transparency, and equity implications. AI systems must be designed and implemented with explicit attention to fairness, avoiding algorithmic bias that might disadvantage particular student populations. The framework emphasizes the importance of maintaining human agency in educational decision-making.

5. Conclusions

The AI-enhanced APIA framework presented in this paper addresses fragmentation in current AI applications across high-stakes examination systems by proposing systematic integration across preparation, implementation, and analysis phases. The model attempts to bridge the current gaps by synthesizing AI-powered student preparation, predictive modelling, and post-examination analysis within a coherent pedagogical structure that maintains human experience and pedagogical expertise.

The framework's contribution lies in its three-dimensional approach that addresses processual phases, stakeholder responsibilities, and cognitive-taxonomic alignment simultaneously. This structure provides educators and administrators with guidance for implementing AI technologies while avoiding common pitfalls such as technological determinism or cognitive clustering at basic and advanced levels. The emphasis on data-informed decision-making principles ensures that AI applications serve educational objectives rather than constraining pedagogical practices. However, the framework represents an initial theoretical contribution that requires empirical validation across diverse educational contexts.

The framework's relevance extends beyond theoretical contribution to address practical implementation challenges facing educational institutions worldwide. Educational systems seeking to integrate AI technologies can utilize APIA as a structured roadmap for coordinated implementation across multiple stakeholder groups, preventing the misaligned roles and technological disconnects

that have characterized previous AI adoption attempts. The framework's explicit mapping of stakeholder responsibilities and cognitive-taxonomic considerations provides practical guidance for resource allocation, professional development planning, and technology procurement decisions.

The framework's practical value lies in providing educational leaders with a comprehensive perspective and with a structured diagnostic and planning tools for understanding and addressing the intricate relationship between technological capabilities and pedagogical needs.

Note

In the process of drafting the text, AI digital assistants were consulted (ResearchRabbit, Elicit, Sharly, ChatPDF, ChatGPT, Claude) to find additional sources of information, for additions, to improve coherence, clarity and accuracy. All AI suggestions were evaluated and selected by the authors, who fully own the ideas and intellectual contribution in this text.

References

- Akçapınar, G., Altun, A., & Aşkar, P. (2019). Using learning analytics to develop early-warning system for at-risk students. *Int J Educ Technol High Educ, 16,* 40. https://doi.org/10.1186/s41239-019-0172-z
- Anjum, G., Choubey, J., Kushwaha, S., & Patkar, V. (2023). AI in Education: Evaluating the Efficacy and Fairness of Automated Grading Systems. *IJIRSET, 12*(6), 9043–9050. https://doi.org/10.15680/IJIRSET.2023.1206161
- Brazauskienė, E. (2025). Educational Decision-Making in Digital Education: A Conceptual Review of Data-Driven, Data-Based, and Data-Informed Approaches. *Revista de Pedagogie Digitală, 4*(1), 63–72.
- Bulut, O., Beiting-Parrish, M., Casabianca, J. M., Slater, S. C., Jiao, H., Song, D., Ormerod, C. M., Fabiyi, D. G., Ivan, R., Walsh, C., Rios, O., Wilson, J., Yildirim-Erbasli, S. N., Wongvorachan, T., Liu, J. X., Tan, B., & Morilova, P. (2024). The Rise of Artificial Intelligence in Educational Measurement: Opportunities and Ethical Challenges. *Journal of Educational Measurement and Evaluation, 5*(3). https://doi.org/10.48550/arXiv.2406.18900
- Caspari-Sadeghi, S. (2022). Artificial Intelligence in Technology-Enhanced Assessment: A Survey of Machine Learning. *Journal of Educational Technology Systems, 51*(3), 372–386. https://doi.org/10.1177/00472395221138791

- Costa, D., Quinteiro, P., & Dias, A. C. (2019). A systematic review of life cycle sustainability assessment: Current state, methodological challenges, and implementation issues. *Science of The Total Environment, 686,* 774–787. https://doi.org/10.1016/j.scitotenv.2019.05.435
- Cramer, E. D., Little, M. E., & McHatton, P. A. (2014). Demystifying the Data-Based Decision-Making Process. *Action in Teacher Education, 36*(5–6), 389–400. https://doi.org/10.1080/01626620.2014.977690
- Decabooter, I., Warmoes, A., Consuegra, E., Van Gasse, R., & Struyven, K. (2024). The data coach chronicles: A systematic review of human support for making data-based decision-making a success. *Teaching and Teacher Education, 146,* 104641. https://doi.org/10.1016/j.tate.2024.104641
- Du, X., Dû, M. L., Zhou, Z., & Bai, Y.-M. (2025). Facilitator or hindrance? The impact of AI on university students' higher-order thinking skills in complex problem solving. *International Journal of Educational Technology in Higher Education, 22*(1). https://doi.org/10.1186/s41239-025-00534-0
- Fartuşnic, R., Istrate, O., & Fartuşnic, C. (2025). Beyond Automation: A Conceptual Framework for AI in Educational Assessment. *Journal of Digital Pedagogy, 4*(1), 83–102. https://doi.org/10.61071/JDP.2555
- Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: a consolidated model of a field of research and practice. *Learning Research and Practice, 3*(1), 63–78. https://doi.org/10.1080/23735082.2017.1286142
- Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making.* Institute of Education Sciences. https://files.eric.ed.gov/fulltext/ED506645.pdf
- Hartmann, P. (2023). *Digitization of high-stakes exams: Empirical insights and design recommendations for the digital execution and scoring of exams* [Doctoral dissertation, Georg-August-Universität Göttingen]. Göttingen. https://ediss.uni-goettingen.de/handle/11858/14861
- Jose, B., Cherian, J., Verghis, A. M., Varghise, S. M., Mumthas, S., & Joseph, S. (2025). The cognitive paradox of AI in education: between enhancement and erosion. *Frontiers in Psychology, 16*. https://doi.org/10.3389/fpsyg.2025.1550621
- Kostopoulos, G., & Kotsiantis, S. (2022). Exploiting Semi-supervised Learning in the Education Field: A Critical Survey. In G. A. Tsihrintzis, M. Virvou, & L. C. Jain (Eds.), *Advances in Machine Learning/Deep Learning-based Technologies. Learning and Analytics in Intelligent Systems* (Vol. 23, pp. 79–94). Springer, Cham. https://doi.org/10.1007/978-3-030-76794-5_5
- Lahza, H., Smith, T. G., & Khosravi, H. (2023). Beyond item analysis: Connecting student behaviour and performance using e assessment logs. *British Journal of*

Educational Technology, 54(1), 335–354. https://doi.org/10.1111/bjet.13270

- Lee, J., & Lee, J. (2025). Generative AI chatbot for teachers' data-informed decision-making: Effects and insights. *Educational Technology & Society, 28*(3), 298–317. https://doi.org/10.30191/ETS.202507_28(3).TP06
- Lubbe, A., Marais, E., & Kruger, D. (2025). Cultivating independent thinkers: The triad of artificial intelligence, Bloom's taxonomy and critical thinking in assessment pedagogy. *Education and Information Technologies, 30*, 17589–17622. https://doi.org/10.1007/s10639-025-13476-x
- Mandinach, E. B. (2012). A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice. *Educational Psychologist, 47*, 71–85.
- Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(8), em2307. https://doi.org/10.29333/ejmste/13428
- Richardson, M., & Clesham, R. (2021). Rise of the machines? The evolving role of AI technologies in high-stakes assessment. *London Review of Education, 19*(1), 1–13. https://doi.org/10.14324/LRE.19.1.09
- Stokkink, P. (2025). *The Impact of AI on Educational Assessment: A Framework for Constructive Alignment.* https://doi.org/10.48550/ARXIV.2506.23815
- Thompson, G., & Sellar, S. (2018). Datafication, testing events and the outside of thought. *Learning, Media and Technology, 43*(2), 139–151. https://doi.org/10.1080/17439884.2018.1444637
- Torrisi-Steele, G., Wang, V., & Galaige, J. (2019). Practice Perspectives on Learning Analytics in Higher Education. In V. Wang (Ed.), *Handbook of Research on Transdisciplinary Knowledge Generation* (pp. 45–56). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-5225-9531-1.ch004
- Van der Kleij, F. M., Vermeulen, J. A., Schildkamp, K., & Eggen, T. J. H. M. (2015). Integrating data-based decision making, Assessment for Learning and diagnostic testing in formative assessment. *Assessment in Education: Principles, Policy & Practice, 22*(3), 324–343. https://doi.org/10.1080/0969594X.2014.999024
- Wang, Y. (2021). When artificial intelligence meets educational leaders' data-informed decision-making: A cautionary tale. *Studies in Educational Evaluation, 69*, 100872. https://doi.org/10.1016/j.stueduc.2020.100872
- Williams, P. (2023). AI, analytics and a new assessment model for universities. *Education Sciences, 13*(10), 1040. https://doi.org/10.3390/educsci13101040
- Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9*, 11. https://doi.org/10.1186/s40561-022-00192-z