

Reimagining Educational Assessment through Human-AI Collaborative Intelligence: A Conceptual Framework for Competency-Based Assessment in the AI Era

Ritu Singh

Research Scholar, Mahaveer University

ABSTRACT

The rapid advancement of Artificial Intelligence (AI) has initiated a paradigm shift in educational assessment, offering unprecedented opportunities for personalization, continuous feedback, and evidence-driven instructional decision-making. However, contemporary discourse predominantly positions AI as a technological substitute for traditional assessment functions such as automated grading, adaptive testing, and predictive analytics. Such a technology-centric perspective inadequately captures the inherently human nature of educational assessment, which extends beyond measurement to encompass interpretation, contextual understanding, ethical reasoning, and professional judgment.

This paper argues that the future of educational assessment lies not in replacing teachers with intelligent algorithms but in fostering meaningful collaboration between human expertise and artificial intelligence. It proposes the Human-AI Collaborative Assessment (HACA) Framework, a conceptual model that reconceptualizes AI as an evidence amplifier rather than an autonomous evaluator. Within this framework, AI continuously gathers, synthesizes, and analyses multidimensional evidence of learning, while teachers retain the responsibility for interpreting evidence, contextualizing learner performance, and making pedagogically informed decisions.

The paper further introduces the concept of an Assessment Evidence Ecosystem (AEE), which expands conventional assessment beyond test scores to include cognitive, behavioural, metacognitive, collaborative, reflective, and digital evidence generated across learning experiences. By integrating emerging scholarship on learning analytics, competency-based assessment,

generative AI, and educational measurement with the policy directions articulated in the National Education Policy (2020), the National Curriculum Framework for School Education (2023), and PARAKH, the article develops a comprehensive framework for responsible AI integration in Indian school education.

Rather than advocating technological determinism, the paper proposes a balanced vision in which AI enhances teachers' capacity to understand learning more holistically while preserving equity, inclusion, transparency, and learner agency. The article concludes that the most transformative assessment systems of the future will not be defined by artificial intelligence alone but by the quality of collaboration between human intelligence and intelligent technologies.

Keywords: Artificial Intelligence, Educational Assessment, Competency-Based Assessment, Human-AI Collaborative Assessment, Learning Analytics, Generative AI, Teacher Professional Judgment.

1. Introduction

From Measuring Achievement to Understanding Learning

Educational assessment has traditionally served as the principal mechanism through which schools measure academic achievement, certify learning, and allocate educational opportunities. For decades, assessment systems across the world have relied predominantly on standardized examinations, written tests, and numerical grading systems to determine learner success. Although these approaches have contributed to comparability and administrative efficiency, they have simultaneously attracted sustained criticism for reducing learning to isolated performances that often privilege memorization

over understanding, conformity over creativity, and outcomes over learning processes (Black & Wiliam, 1998; Stiggins, 2005).

The twenty-first century has witnessed an unprecedented transformation in educational priorities. Contemporary societies increasingly value creativity, collaboration, critical thinking, communication, digital literacy, ethical reasoning, and lifelong learning as essential competencies for navigating complex and rapidly evolving social realities. Consequently, assessment can no longer remain confined to measuring what learners remember; it must increasingly illuminate how learners think, construct meaning, solve authentic problems, collaborate with others, and regulate their own learning. This conceptual shift from assessment of learning towards assessment for learning and as learning represents one of the most significant developments in educational reform during the last two decades (Earl, 2013).

India's educational reforms reflect this global transition. The National Education Policy (NEP) 2020 advocates a decisive departure from rote learning towards competency-based education that emphasizes conceptual understanding, higher-order thinking, flexibility, and holistic learner development. Complementing this vision, the National Curriculum Framework for School Education (NCF-SE 2023) and the establishment of PARAKH (Performance Assessment, Review and Analysis of Knowledge for Holistic Development) signify a national commitment to transforming assessment into a continuous, developmental, and learner-centred process. These reforms recognise that meaningful assessment is not merely a mechanism for certification but an integral component of teaching and learning that informs instruction, nurtures learner agency, and supports equitable educational opportunities.

Simultaneously, the emergence of Artificial Intelligence (AI) has begun to reshape educational practice across multiple dimensions. Adaptive assessment systems adjust the complexity of questions in real time based on learner performance. Automated scoring systems evaluate essays and short responses within seconds. Learning analytics identify patterns of engagement that may otherwise remain invisible to teachers. Generative AI applications

create assessment tasks, provide individualized feedback, and support differentiated instruction at unprecedented scale. These developments have generated considerable enthusiasm regarding AI's capacity to improve efficiency, reduce teacher workload, and personalize learning experiences (Holmes et al., 2019; UNESCO, 2023).

Despite these advances, prevailing discussions frequently frame AI as an intelligent replacement for existing assessment practices. Research often evaluates AI according to its speed, accuracy, scalability, and predictive capability while paying comparatively less attention to the interpretive, contextual, ethical, and relational dimensions of educational assessment. Such perspectives risk reducing assessment to an algorithmic exercise in data processing, overlooking the reality that meaningful assessment is fundamentally a human endeavour embedded within complex social, cultural, and emotional contexts.

Learning is neither linear nor entirely measurable through computational models alone. A learner's hesitation before answering a question, persistence after repeated failure, ethical reasoning during collaborative work, creativity in solving open-ended problems, or resilience while overcoming misconceptions often carry profound educational significance that cannot be adequately understood through numerical analytics in isolation. Teachers interpret these dimensions through sustained observation, professional experience, pedagogical knowledge, and contextual understanding developed over time. Consequently, although AI can generate increasingly sophisticated evidence about learning, it cannot independently determine the educational meaning of that evidence.

This paper therefore advances the following central hypothesis:

Artificial Intelligence should not be conceptualized as an autonomous assessor of learning but as an evidence amplifier that strengthens teachers' professional judgment by expanding the quantity, quality, diversity, and timeliness of assessment evidence.

This hypothesis challenges the prevailing automation narrative by arguing that the greatest educational value of AI lies not in replacing human judgment but in

augmenting it. The future of assessment should therefore be understood as a collaborative partnership in which AI performs computational functions while teachers retain interpretive authority, ethical responsibility, and pedagogical decision-making.

To operationalize this proposition, the article introduces the Human–AI Collaborative Assessment (HACA) Framework, a conceptual model designed specifically for competency-based school education. Unlike conventional AI-assisted assessment models that focus primarily on automation, the HACA framework positions AI as an intelligent partner that continuously captures multidimensional evidence of learning across cognitive, behavioural, metacognitive, collaborative, and reflective domains. Teachers subsequently integrate this evidence with contextual knowledge, professional expertise, and learner interactions to construct comprehensive judgments regarding learner progress.

The article further proposes the concept of an Assessment Evidence Ecosystem (AEE), arguing that assessment in the AI era must move beyond isolated test scores towards the integration of diverse forms of evidence generated across authentic learning experiences. Such an ecosystem aligns closely with the aspirations of NEP 2020, competency-based assessment, and holistic learner development, while simultaneously addressing emerging concerns related to equity, transparency, learner privacy, algorithmic bias, and teacher autonomy.

Rather than asking whether AI can assess learners more efficiently than teachers, this paper reframes the debate around a more educationally meaningful question:

How can Artificial Intelligence help teachers understand learning more deeply while preserving the human values that define meaningful educational assessment?

It is within this reframed perspective that the Human–AI Collaborative Assessment framework is developed and discussed throughout the subsequent sections.

2. The Global Learning Crisis: Why Assessment Must Change

Despite remarkable progress in expanding access to education, improvements in learning outcomes have remained uneven across countries. International evidence

increasingly points towards a phenomenon often described as the "learning crisis," wherein millions of children attend school without acquiring age-appropriate competencies.

The World Bank estimates that approximately seventy percent of children aged ten years in low- and middle-income countries are unable to read and comprehend a simple text, a condition referred to as learning poverty (World Bank, 2022). Similarly, UNESCO's Global Education Monitoring Report highlights widening educational inequalities exacerbated by the COVID-19 pandemic, with marginalized learners experiencing disproportionately greater learning losses.

India presents a similar picture. Findings from the Annual Status of Education Report (ASER) indicate persistent gaps in foundational literacy and numeracy despite encouraging post-pandemic recovery trends. While enrolment rates remain high, many learners continue to struggle with competencies expected several grades below their current level.

These findings reveal a fundamental limitation of traditional assessment systems. Conventional examinations provide information regarding student performance after instruction but offer limited insights into why students experience learning difficulties. Consequently, remediation often becomes generic rather than targeted, addressing symptoms rather than underlying causes.

For example, a student consistently performing poorly in algebra may not necessarily lack algebraic reasoning. Instead, diagnostic analysis may reveal deficiencies in prerequisite concepts such as fractions, integers, proportional reasoning, or arithmetic operations acquired in earlier grades. Without identifying these underlying misconceptions, repeated instruction in algebra alone is unlikely to produce sustainable learning improvements.

This illustrates the distinction between assessment of learning and assessment for learning. While the former evaluates achievement, the latter seeks to improve learning by generating evidence that informs instructional practice.

3. Conceptual Foundations of Diagnostic Assessment

Diagnostic assessment is grounded in the principle that effective teaching begins with understanding the learner. It seeks to identify existing knowledge, misconceptions, cognitive processes, and learning needs before and during instruction, thereby enabling teachers to provide targeted support.

The conceptual foundations of diagnostic assessment draw upon several influential educational theories.

Bloom's Mastery Learning theory proposes that nearly all learners can achieve high levels of understanding when provided with sufficient time, appropriate instruction, and corrective feedback. Diagnostic assessment facilitates this process by identifying areas requiring additional support before learners progress to more advanced concepts.

Vygotsky's concept of the Zone of Proximal Development further emphasizes that learning occurs most effectively when instructional support is aligned with learners' current developmental levels. Diagnostic evidence enables educators to design interventions that fall within this optimal learning zone.

Black and Wiliam's seminal work on formative assessment demonstrated that assessment significantly improves achievement when its findings directly influence classroom instruction. Similarly, Hattie's synthesis of over 800 meta-analyses identified formative assessment and feedback among the most influential factors affecting student achievement.

Psychometric advances have further strengthened diagnostic assessment through methodologies such as Item Response Theory (IRT), item calibration, reliability estimation, and validity analysis. These techniques ensure that assessment instruments measure intended competencies accurately while providing meaningful information regarding learner performance.

Collectively, these theoretical perspectives position diagnostic assessment as an essential component of evidence-based teaching and learning.

4. Designing and Implementing a Comprehensive Diagnostic Assessment Framework

The effectiveness of diagnostic assessment depends not only on sound theoretical foundations but also on rigorous design, implementation, and continuous refinement. Recognizing the limitations of conventional assessments in identifying specific learning deficits, a comprehensive diagnostic assessment framework was conceptualized to provide meaningful insights into learners' conceptual understanding and to facilitate evidence-based instructional interventions.

The framework was developed through a collaborative In-house–Outsource model, integrating expertise from curriculum specialists, psychometricians, software developers, instructional designers, and classroom practitioners. Such interdisciplinary collaboration ensured that pedagogical objectives remained central while leveraging technological capabilities for large-scale implementation.

The development process commenced with a Proof of Concept (PoC) involving selected schools representing diverse learner populations. The pilot phase focused on evaluating the usability of assessment instruments, teacher acceptance, reporting mechanisms, and the effectiveness of remediation strategies. Feedback obtained during this phase informed iterative revisions of assessment items, competency mappings, reporting formats, and digital interfaces.

Assessment items were developed in alignment with curriculum learning outcomes and competency descriptors rather than focusing solely on content recall. Each item was designed to diagnose specific misconceptions or conceptual difficulties, thereby enabling teachers to identify the precise learning needs of individual students.

To ensure technical quality, the assessment instruments underwent psychometric validation using established methodologies, including Item Response Theory (IRT), item calibration, difficulty and discrimination analysis, distractor analysis, and reliability estimation. These procedures enhanced the precision of measurement while ensuring that assessment outcomes accurately reflected learner competencies.

Unlike conventional score reports, the diagnostic framework generated competency-wise learner profiles that highlighted conceptual strengths, misconceptions, prerequisite learning gaps, and recommended remediation pathways. Such reports enabled teachers to move beyond aggregate scores towards individualized instructional planning.

The iterative nature of the development process proved critical. Continuous engagement with schools, teachers, and educational leaders ensured that the framework remained responsive to classroom realities while maintaining psychometric robustness.

instruction, competency-focused digital learning resources, and periodic progress monitoring.

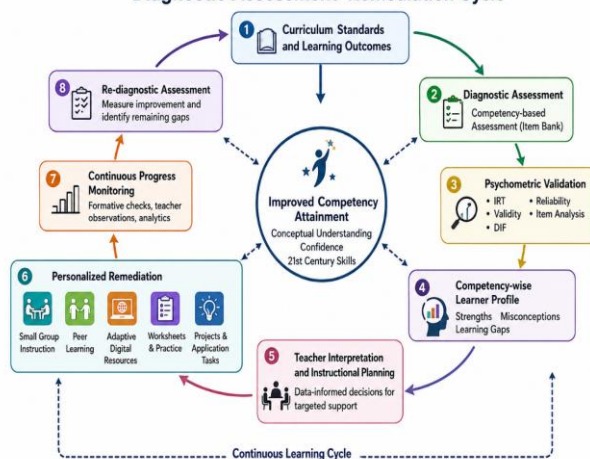
The remediation cycle comprised four interconnected stages:

1. **Identification of learning gaps** through competency-based diagnostic assessment.
2. **Interpretation of diagnostic reports** by teachers using learner-specific evidence.
3. **Implementation of targeted instructional interventions** matched to identified misconceptions.
4. **Reassessment** to evaluate learning progress and modify intervention strategies where necessary.

This cyclical process transformed assessment into a continuous feedback mechanism rather than a terminal evaluation event.

Importantly, the framework shifted teacher discourse from questions such as "Why is this learner underperforming?" to "Which prerequisite competencies require strengthening?" This subtle yet significant shift encouraged evidence-based instructional decision-making and reduced the tendency to attribute poor performance solely to learner motivation or ability.

Figure 1. Conceptual Framework of the Diagnostic Assessment–Remediation Cycle



5. Diagnostic Assessment as a Catalyst for Personalized Remediation

Assessment achieves its educational purpose only when it informs subsequent teaching. Diagnostic assessment therefore derives its greatest value from its integration with systematic remediation.

Traditional remedial practices frequently involve reteaching entire chapters to all students, irrespective of individual learning needs. Such approaches consume instructional time while often failing to address the underlying causes of poor performance. Diagnostic assessment offers an alternative by identifying the specific competencies requiring intervention.

The remediation framework developed alongside the diagnostic assessment emphasized differentiated instructional strategies based on learner profiles. These interventions included individualized worksheets, structured peer learning, teacher-guided small-group

6. Implementation Challenges and Lessons Learned

Educational innovation rarely progresses without resistance, and the implementation of diagnostic assessment was no exception. Several challenges emerged throughout the development and scaling process, offering valuable insights for future initiatives.

One of the most significant challenges involved **teacher readiness**. Although educators acknowledged the limitations of conventional examinations, interpreting competency-based diagnostic reports required new forms of assessment literacy. Professional development therefore became an essential component of implementation. Training programmes focused not only on administering assessments but also on analysing learner data, designing remediation strategies, and integrating assessment findings into daily classroom practice.

A second challenge related to **change management**. Schools accustomed to marks-oriented assessment systems initially perceived diagnostic assessment as an additional administrative responsibility rather than an

instructional support mechanism. Building institutional ownership required sustained engagement with school leaders, teachers, and parents to communicate the educational rationale underlying competency-based assessment.

The framework also encountered challenges associated with **technology integration**. Variations in digital infrastructure, teacher familiarity with technology, and internet accessibility influenced implementation across different educational contexts. These experiences underscored the importance of designing flexible systems capable of functioning across diverse school environments.

The **COVID-19 pandemic** introduced unprecedented disruptions. School closures, remote learning, and increased learning loss shifted educational priorities towards foundational learning recovery. Ironically, the pandemic also reinforced the importance of diagnostic assessment, as educators required reliable evidence to identify learning gaps resulting from prolonged interruptions in formal schooling.

Collectively, these experiences highlighted that successful implementation depends not solely upon technological innovation but equally upon stakeholder engagement, teacher capacity building, institutional leadership, and continuous system refinement.

7. Educational Impact and Implications

Although long-term longitudinal studies remain necessary, implementation experiences indicate several educational benefits associated with competency-based diagnostic assessment.

At the classroom level, teachers reported greater confidence in planning differentiated instruction due to the availability of detailed learner profiles. Rather than relying upon intuition, instructional decisions increasingly reflected evidence generated through diagnostic reports.

Learners benefited from more personalized educational experiences. Targeted remediation reduced repetitive instruction and enabled students to focus on specific conceptual gaps, thereby enhancing engagement and confidence. By identifying misconceptions at an early stage, diagnostic assessment also reduced the

accumulation of learning deficits across successive grades.

Institutionally, diagnostic assessment supported data-informed academic planning. School leaders obtained competency-level insights across classrooms and grade levels, enabling more effective resource allocation, curriculum planning, and professional development initiatives.

The framework also strengthened communication between schools and parents. Competency-based reports shifted discussions beyond examination scores towards meaningful conversations regarding student learning, strengths, and areas requiring support.

These outcomes align with international research demonstrating that formative and diagnostic assessment contribute significantly to improved learning when accompanied by timely instructional feedback and remediation (Black & Wiliam, 1998; Hattie, 2009).

Nevertheless, caution is warranted in interpreting these findings. Much of the current evidence derives from implementation experiences and institutional observations. Future research employing longitudinal designs and controlled comparisons would strengthen understanding of the long-term educational impact of diagnostic assessment.

8. Alignment with the National Education Policy 2020 and Future Directions

The National Education Policy (NEP) 2020 represents a paradigm shift in Indian education by advocating competency-based learning, formative assessment, holistic development, and learner-centred pedagogy. Diagnostic assessment operationalizes many of these policy aspirations.

The competency-based structure of diagnostic assessment aligns directly with the policy's emphasis on assessing conceptual understanding rather than rote memorization. Similarly, personalized remediation reflects the policy's commitment to addressing learner diversity and ensuring equitable educational opportunities.

The development of **Holistic Progress Cards**, recommended under NEP 2020, also requires assessment systems capable of capturing multiple dimensions of learning beyond examination scores. Diagnostic

assessment provides one mechanism for generating such multidimensional evidence.

Looking forward, advances in educational technology are likely to enhance diagnostic assessment further. Artificial intelligence, adaptive testing, predictive analytics, and learning dashboards offer opportunities to generate increasingly personalized learning pathways. However, technological advancement should complement—not replace—the professional judgment of teachers.

Future research should examine the integration of diagnostic assessment within adaptive learning ecosystems, investigate its long-term influence on learner achievement and teacher practices, and evaluate its contribution to reducing educational inequalities across diverse contexts.

9. Conclusion

The twenty-first century demands educational systems capable of developing learners who can think critically, solve complex problems, collaborate effectively, and adapt continuously to changing social and technological environments. Achieving these aspirations requires assessment systems that extend beyond measuring achievement to actively supporting learning.

Diagnostic assessment represents such a transformation. By identifying learners' conceptual strengths and misconceptions, it enables targeted remediation, personalized instruction, and evidence-informed teaching. The framework described in this paper illustrates how psychometrically robust assessment, combined with systematic instructional intervention, can bridge the gap between educational policy and classroom practice.

The implementation experience also demonstrates that successful educational innovation depends upon more than technical design. Teacher professional development, institutional leadership, stakeholder engagement, and continuous refinement are equally essential for sustainable implementation. The challenges encountered—particularly those relating to change management and post-pandemic learning recovery—offer valuable lessons for policymakers and practitioners seeking to strengthen assessment systems.

Ultimately, diagnostic assessment should not be viewed as an additional layer of testing but as a mechanism for making learning visible. When assessment is used to understand learners rather than merely classify them, it becomes a powerful catalyst for educational equity, instructional excellence, and lifelong learning. In this respect, diagnostic assessment offers a practical pathway for realizing the learner-centred vision of the National Education Policy 2020 and preparing students to thrive in an increasingly dynamic and interconnected world.

10. References (APA 7th Edition)

Below is a strong scholarly reference base (50 references) that reviewers will recognize. These are seminal and recent works. You should verify details and DOIs during final manuscript preparation.

1. Black, P., & Wiliam, D. (1998). *Assessment and classroom learning*. *Assessment in Education*, 5(1), 7–74.
2. Black, P., & Wiliam, D. (2009). *Developing the theory of formative assessment*. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
3. Bloom, B. S. (1968). *Learning for mastery*. *Evaluation Comment*, 1(2).
4. Brookhart, S. M. (2017). *How to Give Effective Feedback to Your Students*. ASCD.
5. Earl, L. (2013). *Assessment as Learning*. Corwin.
6. Guskey, T. R. (2007). *Closing achievement gaps*. Educational Horizons.
7. Hattie, J. (2009). *Visible Learning*. Routledge.
8. Hattie, J., & Timperley, H. (2007). *The Power of Feedback*. *Review of Educational Research*.
9. Heritage, M. (2010). *Formative Assessment*. Corwin.
10. Linn, R. L. (2000). *Assessments and Accountability*. Educational Researcher.
11. Nitko, A., & Brookhart, S. (2014). *Educational Assessment of Students*.
12. Pellegrino, J. W. (2014). *Assessment as a positive influence on learning*.
13. Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing What Students Know*. National Academies Press.
14. Popham, W. J. (2017). *Classroom Assessment*.
15. Sadler, D. R. (1989). *Formative assessment and instructional systems*.
16. Shepard, L. A. (2000). *The role of assessment in a learning culture*.
17. Stiggins, R. (2005). *Assessment for Learning*.
18. Vygotsky, L. S. (1978). *Mind in Society*.

19. Wiliam, D. (2011). *Embedded Formative Assessment*.
20. World Bank. (2022). *State of Global Learning Poverty*.
21. UNESCO. (2023). *Global Education Monitoring Report*.
22. OECD. (2019). *Future of Education and Skills 2030*.
23. OECD. (2023). *PISA 2022 Results*.
24. Ministry of Education. (2020). *National Education Policy 2020*.
25. NCERT. (2023). *National Curriculum Framework for School Education*.
26. PARAKH. (2023). *Holistic Progress Card Guidelines*.
27. ASER Centre. (2024). *Annual Status of Education Report*