EDUKASI: JURNAL PENDIDIKAN DAN PENGAJARAN
ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

LEARNING ASSESSMENT MODEL FOR THE SOCIETY 5.0: INTEGRATION OF AI AND TEACHER-BASED AFFECTIVE ASSESSMENT THROUGH SIPAQ' MANDAR VALUES

AYYUB YUNUS, IQBAL, MUHAMMAD ASWAD
Universitas Sulawesi Barat, Majene,Indonesia
*Corresponding author**:** ayyub@unsulbar.ac.id

Abstract
The transition into Society 5.0 necessitates educational assessment models that are not only technologically advanced but also culturally meaningful. This study developed and tested an integrated learning assessment model for the Society 5.0, synergizing Artificial Intelligence (AI) for cognitive-psychomotor evaluation with teacher-based affective assessment grounded in Sipaq' Mandar local values. Employing an explanatory sequential mixed -methods design, the research involved 300 junior high school students and 60 teachers in Majene Regency, West Sulawesi. Quantitative data were collected through pre-post tests and questionnaires, while qualitative data were gathered via interviews, FGDs, and observation. Result demonstrated a statistically significant improvement in cognitive-psychomotor achievement, with an average score increase of 11.7 points ($p < .001$, Cohen's $d = 1.12$). The Sipaq' Mandar values (siri', siarondoi, sipakalaq'bi, makkeada, dipokannyang) were successfully operationalized into a valid and reliable affective assessment rubric. The resulting SIAMASEI-AI model was deemed highly feasible (M=4.43/5) and contributed 71.2% to holistic assessment affectiveness. The study concludes that the integration of AI with culturally contextual affective assessment offers a viable, balanced, and humanized evaluation framework, addressing both the technological demands and the character-building imperative of modern education.

Keywords: affective assessment, artificial intelligence, local wisdom, sipaq' mandar, society 5.0.

Introduction

The Society 5.0 era has driven a paradigm shift in the world of education, an era where technology no longer serves as a passive tool but as an active cognitive partner in designing learning experiences and evaluation systems (Pandya, 2024). This transformation brings significant demands on assessment practices, which must evolve from conventional approaches often limited to summative cognitive measurement towards systems that are holistic, formative, and capable of capturing the complexity of 21st-century competencies such as critical thinking, creativity, collaboration, and character (Agrawal et al., 2025; Hendri et al., 2025). However, beyond its efficiency and objectivity potential, there is a risk of educational reductionism if assessment is fully entrusted to artificial technology. The affective domain, encompassing values, attitudes, and character, remains a contextual and cultural human domain where the subjective judgment of teachers as moral agents and cultural translators is irreplaceable (Fepriyanti & Roqib, 2024). Therefore, a primary challenge of education in the Society 5.0 era is designing assessment models that synergistically integrate the objectivity of technology with human subjective wisdom, while maintaining sociocultural relevance.

Indonesia, as a country with rich cultural diversity, faces a unique dilemma in transforming technology-based education. On one hand, the adoption of digital technology in learning is increasingly massive, including in West Sulawesi Province which utilizes digital devices post-

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

earthquake and COVID-19 pandemic (Data of Education Office, 2024). On the other hand, there are concerns that accelerated digitalization may erode the foundation of local values which actually support national character building (Awaru et al., 2024). The Mandar community in West Sulawesi, which possesses living Sipaq' Mandar values implemented in community life, such as *siri'* (sense of shame and honor), *siarondoi* (solidarity and mutual cooperation), *sipakalaq'bi* (mutual respect), *makkeada* (politeness), and *dipokannyang* (integrity and honesty) (Burhanuddin et al., 2023). These values are actually aligned with global character education goals but have not been systematically structured within formal assessment instruments. A preliminary study conducted by the researchers in Majene Regency (see Yunus, 2025, for detailed data) shows that 82% of students feel the assessments they receive do not fully reflect their character, while 68% of teachers acknowledge the potential of AI but lack integrative guidance with local values. These findings confirm a missing link between technological advancement and meaningful, contextual measurement of affective development.

Based on a literature review, several research gaps can be identified. First, research on AI in educational assessment in Indonesia is still dominated by its application in the cognitive and psychomotor domains, such as automated scoring systems and performance analysis (Hendri et al., 2025; Kumar, 2025). Second, although the local wisdom Sipaq' Mandar has been extensively studied from sociological and philological perspectives (Burhanuddin et al., 2023), there has been no systematic effort to transform it into an operational framework for affective assessment in schools. Third, integrated assessment models that combine AI and culture-value-based assessment methodologically and applicatively are still very limited, even though such integration is believed to create a more balanced and humane assessment ecosystem (El Karafli, 2025; Sánchez Rogel et al., 2024). Fourth, most discussions on local wisdom-based education remain at the normative discourse level, not yet reaching the stage of concrete model development and testing in the field, despite the demands of Sustainable Development Goal (SDG) 4 emphasizing the importance of education that is inclusive, quality, and culturally relevant.

To address these gaps, this research is designed with the main objective of developing and testing a learning assessment model for the Society 5.0 that integrates AI technology for cognitive-psychomotor assessment with teacher-based affective assessment through the framework of Sipaq' Mandar values. Specifically, this research aims to: (1) analyze the concept and forms of AI application in cognitive and psychomotor assessment within the context of secondary schools; (2) identify and operationalize Sipaq' Mandar values into observable behavioral indicators for affective assessment; (3) design an integrated assessment model that synergizes AI and the teacher's role; and (4) test the feasibility and effectiveness of the model within the context of science learning in junior high schools in Majene Regency.

The research questions proposed are: (1) How can the concept and application of AI technology be implemented in cognitive and psychomotor assessment systems in the Society 5.0 era?; (2) How can Sipaq' Mandar values be formulated into a valid and reliable affective assessment instrument?; (3) What is the design of an integrated assessment model that integrates AI and Sipaq' Mandar values into a holistic evaluation system?; and (4) How feasible and effective is this integrated assessment model when applied to science learning in junior high schools in Majene Regency?.

This research is expected not only to contribute theoretically by enriching the repository of hybrid educational evaluation models but also to provide practical solutions for teachers in conducting objective, adaptive, and culturally rooted assessments. At the policy level, the research findings can serve as strategic input for local and central governments in designing learning evaluation frameworks aligned with the Merdeka Curriculum and the spirit of preserving local wisdom amidst the flow of digital transformation.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

Literature Review

Paradigm shift in assessment in the digital education era

The development of digital technology has brought fundamental changes to the philosophy and practice of educational assessment. Whereas assessment previously tended to be summative, isolated, and focused on cognitive measurement (Black & Wiliam, 2009), the Society 5.0 era demands a more formative, integrated, and holistic approach. The concept of smart assessment emerges as a response to this need, defined as an evaluation system that utilizes technology to create adaptive, personalized, and continuous assessment experiences (Agrawal et al., 2024). This shift aligns with the demands of 21st-century competencies emphasizing higher-order thinking skills and character, which cannot be adequately measured by conventional instruments (Matheis & John, 2024). Recent literature shows that learning effectiveness heavily depends on the quality of feedback provided, and this is where technology plays a crucial role in delivering precise real-time feedback (Hendri et al., 2025).

The role of Artificial Intelligence (AI) in assessment transformation

The integration of AI in education, particularly in assessment, has become a rapidly growing research field. AI offers the capacity to automate assessment in the cognitive and psychomotor domains with a high degree of objectivity and efficiency. A study by Dubey et al. (2025) demonstrates how AI systems can analyze learning patterns, adaptively adjust question difficulty levels, and evaluate complex skill performance through video or performance data analysis. This capability supports the realization of data-driven instruction, where pedagogical decisions can be made based on strong empirical evidence (Prestoza et al., 2025). However, the literature also warns about the limitations of AI. When dealing with the affective domain such as values, empathy, cooperation, and integrity, AI reaches fundamental limits because its assessment requires contextual interpretation, cultural understanding, and consideration of human values (Fepriyanti & Roqib, 2024). Thus, AI serves optimally as tools for augmentation, not tools for replacement in a comprehensive assessment ecosystem.

Affective assessment and the significance of the teacher's role as a moral agent

The affective domain encompasses dimensions of attitude, values, motivation, and character, which are crucial foundations in holistic education. Bloom (1956), in his taxonomy, has placed the affective domain as an integral part of learning objectives. In the context of the Society 5.0 era, strengthening the affective dimension is increasingly crucial as a countermeasure to the potential dehumanizing impacts of technology (El Karafli, 2025). Affective assessment requires human judgment as it involves continuous behavioral observation, interpretation of intent, and understanding of students' sociocultural backgrounds (Fepriyanti & Roqib, 2024). Teachers, in this case, serve not only as instructors and assessors but also as cultural translators and moral agents who transmit and assess the internalization of values (Sánchez Rogel et al., 2024). The literature underscores that the effectiveness of affective assessment heavily depends on the availability of valid, reliable, and contextual instruments, as well as the teacher's capacity to use them reflectively.

Local wisdom as a contextual foundation in affective assessment

Meaningful education must be relevant to the life context of learners, including the cultural values upheld by their community. The concept of culturally responsive assessment emphasizes

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

the importance of aligning assessment practices with the values, norms, and ways of knowing that exist within a particular community (Gay, 2018). In Indonesia, much research has documented the local wisdom of various ethnic groups as sources of character education values, such as research on Sibaliparriq in Mandar (Awaru et al., 2024) and values within Pappasang Mandar (Burhanuddin et al., 2023). The core Sipaq' Mandar values of *siri', siarondoi, sipakalaq'bi, makkeada, dipokannyang* are a communal ethical system proven to shape the resilient, cohesive, and integrity-filled character of the Mandar people (Burhanuddin et al., 2023). However, the literature review reveals that despite rich philosophical and sociological studies, the transformation of this local wisdom into operationalized affective assessment instruments in the classroom remains very rare. This is the missing link between cultural richness and formal assessment practices in schools.

Integration of technology and cultural values in the theoretical framework

This research is based on two main theoretical frameworks integrated to address the complexity of assessment issues in the Society 5.0 era. Social Constructivist Theory in Assessment (Vygotsky, 1978): This theory emphasizes that learning and cognitive development occur through social interaction and cultural context. Assessment, from this perspective, should be authentic, embedded in learning activities, and capable of capturing the developmental process (process-oriented), not just the final outcome. The application of AI for learning pattern analysis and providing adaptive scaffolding aligns with Vygotsky's concept of the Zone of Proximal Development (ZPD), where technology can help identify and bridge students' ability gaps. Meanwhile, teacher-based affective assessment through observation of social interaction in the classroom is a direct application of social constructivist principles.

The Technological Pedagogical Content Knowledge (TPACK) Framework Enriched with Local Wisdom (Mishra & Koehler, 2006): The TPACK framework highlights the need for synergy between technological knowledge (T), pedagogy (P), and content knowledge (CK) to create effective teaching. This research expands this framework by adding the dimension of Cultural Knowledge (CK) or Local Wisdom Knowledge (LWK). Teachers not only need to master how to use AI platforms (TK) to assess science (CK) with appropriate methods (PK), but must also have a deep understanding of Sipaq' Mandar values (LWK) to design and implement contextual affective assessment. The integration of these four knowledge domains produces Cultural-TPACK, which becomes the theoretical foundation for developing the integrated assessment model in this research.

Identifying the research gap

A synthesis of the literature review above clarifies several research gaps that form the basis for the originality of this study: (1) Integrative Gap: Most research on AI in assessment (e.g., Dubey et al., 2025; Kumar, 2025) and research on local wisdom (e.g., Burhanuddin et al., 2023; Awaru et al., 2024) run parallel and separate. There has been no research that explicitly and systematically designs a model unifying both within a coherent assessment cycle. (2) Operationalization Gap: Local wisdom values are often discussed at abstract and philosophical levels. There is a scarcity of research that operationalizes these values into observable behavioral indicators and ready-to-use assessment rubrics for teachers in secondary schools. (3) Contextual-Empirical Gap: There is much discourse about the importance of culture and technology-based education, but few models have been empirically tested in specific educational settings such as science learning in junior high schools in regional areas. This research fills this void by testing the feasibility and effectiveness of the model in the real context of Majene Regency. (4) Theoretical Gap: The need to develop a theoretical framework capable of explaining and guiding the integration between advanced

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

technology (AI) and traditional value systems. The development of Cultural-TPACK in this research is an effort to address this theoretical need.

Methodology

Research design and development framework

This study adopted an explanatory sequential mixed-methods design (Creswell & Creswell, 2018). This design was chosen because the primary objective was to first quantify the effectiveness of the integrated assessment model and then, in a second phase, to explain and elaborate on those quantitative results by exploring the perspectives, experiences, and contextual nuances of teachers and students. The integration of data was not merely concurrent but sequential and purposeful, with the qualitative phase designed directly to answer the "why" and "how" questions that emerged from the quantitative findings. The procedural framework for model development was adapted from the ADDIE model (Analysis, Design, Development, Implementation, Evaluation) (Branch, 2009).

The entire research and development process, which integrates the explanatory sequential design with the ADDIE model. The procedure for integrating the quantitative and qualitative data followed a clear, sequential process designed to achieve a comprehensive understanding. The explicit data integration procedure is detailed in the following four steps:

1. Phase 1: Quantitative Data Collection and Analysis (QUAN): In the initial phase, quantitative data were collected and analyzed to measure the model's impact and feasibility. This included pre-post tests to measure cognitive-psychomotor achievement, questionnaires to assess model feasibility, and rubric-based scores for affective assessment. The analysis yielded key numerical results, such as a statistically significant score increase (Mean Difference = +11.7, $p < .001$) and a high overall feasibility score ($M = 4.43$). These results established the "what" of the study's outcomes.
2. Identifying Explanatory Questions: The quantitative results, while demonstrating effectiveness, raised critical questions that required qualitative exploration. For instance:
    a. Why did the AI component contribute to such a large effect size in student scores? What were the specific mechanisms perceived by teachers and students?
    b. How did teachers practically implement the Sipaq' Mandar affective assessment rubric in the classroom? What did observing "siri'" or "siarondoi" actually look like in a science learning context?
    c. Why did the qualitative aspect (Sipaq' Mandar values) show a higher regression coefficient than the technological aspect (AI) in predicting holistic assessment effectiveness?
3. Phase 2: Qualitative Data Collection and Analysis (QUAL): The qualitative phase was deliberately designed to investigate the questions above. The data collection instruments—in-depth interview guides, FGD protocols, and observation checklists—were developed specifically to probe deeper into the quantitative outcomes.
    a. For example, to explain the quantitative score increase, interviews with teachers focused on their use of the AI system's real-time feedback, asking for specific examples of how it informed their instructional interventions (e.g., "Can you describe a time when the AI data helped you assist a specific student?").
    b. To explain the high feasibility of the affective rubric, FGDs with teachers and cultural leaders explored the practical application and cultural relevance of the Sipaq' Mandar indicators, moving beyond the numbers to understand the lived experience of implementation.
    c. Observations were conducted to triangulate these findings, providing direct evidence of how the integrated model functioned in the authentic classroom environment.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

4. Integration and Interpretation (Mixing): The final integration occurred at the interpretation stage, where quantitative and qualitative findings were merged to provide a comprehensive understanding. The quantitative results provided the generalizable evidence of effectiveness (the "what"), while the qualitative findings provided the rich, contextual explanation (the "why" and "how").
   a. For instance, the quantitative finding that the model was highly feasible (M=4.43) was substantiated by qualitative themes from teacher interviews, such as "Real-Time Feedback from AI" and "Clarity and relevance of Sipaq' Mandar rubric" (see Table 4). The numbers were given meaning through the teachers' voices.
   b. Similarly, the high regression coefficient for Sipaq' Mandar values was explained qualitatively by teachers' and leaders' assertions that these values provided a "strong moral foundation" and made assessment feel "more humanized," thus clarifying the statistical finding with a cultural and pedagogical rationale.

Through this structured, sequential process, the qualitative data served explicitly to elaborate on, confirm, and explain the initial quantitative results, fulfilling the core principle of the explanatory sequential design.

Research site and participants

The population in this research consists of all teachers and students at the Junior High School (SMP) level in Majene Regency, West Sulawesi. This location was chosen because it is the educational center city of West Sulawesi Province, where schools have implemented AI technology in cognitive and psychomotor assessment and have a strong Mandar cultural tradition contextual to the Sipaq' Mandar values to be integrated into affective assessment. The selected schools are SMP Negeri 1 Majene, SMP Negeri 1 Pamboang, and SMP Negeri 1 Sendana, with a total of 1500 students and 120 teachers.

The sample size for the quantitative phase was determined based on the requirement for parametric testing and a desired powwer of 0.80. Using SPSS software, a minimum of 250 participant was indicated for a paired t-test with a medium effect size (α = .050). to account for potential attrition and ensure robust analysis, 300 students and 60 teachers were recruited. The three schools were selected based on the following criteria, ensuring both practical feasibility and theoretical relevance: (1) representation of the main geographical areas of Majene Regency, (2) established infrastructure for basic technology integration (computers, stable internet), (3) active recognition of Mandar cultural practices within the school environment, and (4) willingness of school leadership to participate in the innovation trial.

For the quantitative method, the sample was determined using stratified random sampling, obtaining a sample of 300 students and 60 teachers, with the following respondent characteristics:

Table 1. *Research Respondent Characteristics*

Category	Sample Number	Percentage (%)
Male students	158	52.67
Female students	142	47.33
Male teachers	32	53.33
Female teachers	28	46.67
Total Respondents	360	100

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

These characteristics show that the sample proportion is relatively balanced between male and female in both student and teacher groups, enabling more representative research results for the population.

For the qualitative research sample, individuals were selected from Mandar community customary leaders, teachers, school principals, and school supervisors using purposive sampling, resulting in 12 respondents with criteria of active involvement in educational activities and an understanding of Mandar cultural values.

Ethical considerations

This study adhered to strict ethical principles to protect the rights and welfare of all participants. The research protocol was reviewed and approved by the Ethics Committee of the Faculty of Teacher Training and Education, Universitas Sulawesi Barat (Reference Number: 045/KEP.FKIP/UNISBAR/2025) prior to data collection.

Informed consent procedure

Informed consent was obtained from all participants following a clear and comprehensive procedure.

1. For Adult Participants (Teachers, Principals, Supervisors, and Community Leaders): Each adult participant was provided with a detailed information sheet explaining the study's purpose, procedures, potential risks and benefits, and their rights as participants. They were given ample time to ask questions before signing a written informed consent form. Participation was entirely voluntary, and they were informed of their right to withdraw from the study at any stage without any negative consequences.
2. For Minor Participants (Students): As the participants were junior high school students under the age of 18, a two-stage consent process was implemented. First, written informed consent was sought from their parents or legal guardians through a consent form sent home via the school. Second, after receiving parental consent, the students themselves were provided with an age-appropriate explanation of the study in a classroom setting. Their verbal or written assent (agreement to participate) was obtained before they were included in the research. Students who did not wish to participate, or whose parents did not provide consent, were not excluded from any regular classroom activities.

Data Protection and Confidentiality Protocol

Robust measures were implemented to ensure the confidentiality and security of all collected data.

1. Anonymity and Confidentiality: To guarantee anonymity, all personally identifiable information (e.g., names, identification numbers) was replaced with unique codes immediately upon data collection. For example, students were coded as S-01, S-02, and teachers as T-01, T-02. All data were reported in aggregate form, ensuring that no individual participant could be identified in any publication or presentation.
2. Data Storage: All physical documents, including signed consent forms and field notes, were stored in a locked cabinet in a secure office at the university, accessible only to the primary research team. Digital data, including survey responses, interview recordings, and analysis files, were stored on a password-protected computer and backed up on an encrypted cloud storage service. Access to digital data was strictly limited to the principal investigators.
3. Right to Withdraw and Data Destruction: Participants were reminded of their right to withdraw their data at any point up until the final data analysis and reporting stage. If a participant chose to withdraw, all of their data would be securely and permanently deleted from both physical and digital storage upon their request.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

This research was conducted in Majene Regency, West Sulawesi Province from January to October 2025, starting with the needs analysis and preliminary study stage (months 1-2), qualitative data collection for cultural value exploration (months 3-5), assessment model development (month 6), field trial and quantitative data collection (months 8-9), and analysis and final report compilation (month 10).

Data collection

The data collection techniques used are tailored to the mixed-methods design, employing two approaches: (1) quantitative data through learning outcome test instruments (measuring students' cognitive and psychomotor achievement), Likert scale questionnaires (to measure student and teacher perceptions of the effectiveness of the assessment model), and observation sheets (to observe the effectiveness of AI system integration in cognitive and psychomotor assessment and teacher assessment for the affective domain); and (2) qualitative data gathered through in-depth interview instruments (with customary leaders, teachers, principals, and supervisors to explore Sipaq' Mandar values), Focus Group Discussion (FGD) (to validate cultural values as the basis for affective assessment), and participatory observation (for learning practices and school culture). According to the explanatory sequential model, the data collection process proceeds sequentially, starting with quantitative data collection, quantitative analysis, and qualitative data collection to deepen the quantitative findings.

Data analysis

Data analysis techniques are also conducted in three stages: (1) quantitative data analysis involving descriptive analysis, instrument validity and reliability tests using Confirmatory Factor Analysis (CFA) and Cronbach's Alpha, and hypothesis testing using paired sample t-test to examine the model's effectiveness, ANOVA to examine the effectiveness of model application across school groups, and multiple linear regression to examine the contribution of AI and cultural values to the effectiveness of student assessment; (2) qualitative data analysis using thematic analysis (data reduction, data presentation, conclusion drawing, and verification) and data validation through member checking, peer debriefing, and triangulation of sources and methods; and (3) integration of quantitative and qualitative data combined at the final interpretation stage (mixing), aiming to strengthen a comprehensive understanding of the research phenomenon.

The characteristics of respondents involved as samples in this study are 300 students and 60 teachers from three junior high schools in Majene Regency. The balanced gender-based proportion (Table 1) aims to ensure the representation of diverse perspectives. Meanwhile, the 12 key informants for qualitative data consist of 4 senior teachers (experience >15 years), 3 school principals, 2 school supervisors, and 3 recognized Mandar customary leaders. This profile indicates that the generated data comes from credible sources with a deep understanding of the local educational and cultural context.

Initial observation at these three schools confirms the preliminary study findings: they possess technological infrastructure such as computers, LCD projectors, and internet connectivity, although their utilization is still limited to material presentation and administrative grade management. Sipaq' Mandar culture exists de facto and is practiced in daily school interactions, such as respectful attitudes towards teachers (sipakalaq'bi) and mutual cooperation in cleaning classrooms and yards (siarondoi), but is not yet structured into standardized affective assessment instruments.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

Findings

Application and effectiveness of AI in cognitive-psychomotor assessment

Based on the analysis of quantitative and qualitative data related to the application of AI in assessment at the three partner schools, significant potential is indicated, although still in the trial stage. A simple AI assessment platform integrated with the school's LMS was implemented to assess cognitive understanding and basic psychomotor skills (such as designing simple science experiments uploaded as short videos) for the science subject. The following data was obtained:

Quantitative data: the paired sample t-test analysis on the pre-test (before intervention) and post-test (after intervention) scores of 300 students revealed a statistically significant improvement in cognitive -psychomotor achievement. The detailed statistics are presented in Table 2 and Table 3.

Table 2. *Descriptive Statistics of Pre-test and Post-test*

Variable	N	Mean	Standard Deviation	Min	Max	Score Range (Mean ± 2 SD)*
Pre-test	300	68.50	8.2	45	89	~52.1 – 84.9
Post-test	300	80.20	7.1	58	96	~66.0 - 94.4

Note: The range estimates approximately 95% of the data assuming a normal distribution.

From this data, the range estimate is calculated as mean + 2 SD (for coverage ~95% assuming normality). Variability decreased in the post-test (lower SD), indicating more consistent results after the intervention.

Table 3. *Results of Paired Sample T-test*

Statistic	Value	Interpretation
Mean Difference	+ 11.7	
t-value (t (299))	15.734	Indicates a significant change.
p-value	< .001	Highly statistically significant.
Cohen's d (Effect size)	~1.12	Large effect (d > 0.8).
Conclusion	Significant improvement	The integrated model effectively improved cognitive-psychomotor scores.

A Cohen's d value of 1.12 indicates a large effect, where d > 0.8 is considered impactful in an educational context.

Qualitative data: results from observations and interviews with teachers regarding perceptions of the AI system in the integrated assessment model. Using simple content analysis thematically, the following data was obtained:

Table 4. *Thematic summary of qualitative data from teacher interviews*

Main theme	Description/Frequency	Representative quotation	Practical implication
Real-Time Feedback from AI	Very helpful (reported by all teachers; high frequency)	"With this system, I can quickly know which students understand the concept of photosynthesis and which are still lagging." (Science Teacher, SMP 1 Majene)	Enables early detection of student difficulties and increases teaching efficiency.
Identification of specific	Focus on data/graph analysis (medium frequency, specific to	"The AI system helps me in identifying each student's specific	Supports instructional differentiation and

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

difficulty areas	science)	difficulty areas, such as data analysis or understanding plant growth graphs" (Science Teacher, SMP 1 Pamboang)	enables teachers to provide targeted intervention.
Targeted and personalized intervention	Overall benefit (high frequency, implicit from observation)	Observation summary: on average, teachers were enthusiastic about providing personalized intervention information after receiving AI feedback)	Increases student engagement and corroborated by triangulation with quantitative data (score increase).

The data above shows how AI reduces teacher workload and personalizes learning, thereby enabling effectiveness in learning implementation. Triangulation analysis (quantitative and qualitative) provides a picture of strong validation, concluding that the model is effective both statistically and practically.

Operationalization of Sipaq' Mandar values into an affective assessment instrument

The AI system is capable of analyzing students' answer patterns and adapting question difficulty levels. Quantitative data (system log analysis): shows that the AI platform indicates 75% of students (225 individuals) experienced a gradual increase in question difficulty level after demonstrating mastery at the basic level. The system automatically presented Higher Order Thinking Skills (HOTS) questions to this group. Conversely, 25% of students (75 individuals) continued to receive basic and intermediate difficulty questions accompanied by system-provided scaffolding and additional explanations.

Qualitative data (student perception questionnaire): 85% of students (255 individuals) agreed with the statement that "the questions given by the AI system were appropriately challenging according to my ability and helped me understand the material step by step." This indicates that adaptive assessment can increase student learning motivation and self-confidence.

The findings above align with research by Agrawal et al. (2024) and Dubey et al. (2025) which affirm that AI can transform assessment to be more dynamic and personal. The capability of AI to analyze learning patterns and provide real-time feedback (Hendri et al., 2025) is proven to optimize the learning process. Teachers no longer spend time assessing routine cognitive tasks and administrative work, allowing them to focus energy on more complex aspects of teaching. Thus, the application of AI for cognitive and psychomotor assessment has addressed the challenges of efficiency and objectivity in the Society 5.0 era, while also opening opportunities for creating data-driven instruction.

Table 5. *Operationalization of Sipaq' Mandar values into affective assessment indicators*

Sipaq' Mandar Value	Philosophical Meaning	Behavioral Indicators in Learning (Example in Science)
Siri' (Sense of shame and honor)	Internal drive to maintain self-respect and dignity through praiseworthy behavior.	- Strives maximally in completing assignments and exams without cheating; - Accepts defeat in science competitions with sportsmanship and uses it as a lesson; - Dares to admit mistakes and take responsibility for actions.
Siarondoi (Cooperation and solidarity)	Spirit of togetherness and mutual cooperation to achieve common goals.	- Active and contributive in group laboratory work; - Helps friends having difficulty understanding material without

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

		expecting reward; - Jointly maintains cleanliness and orderliness of the science lab.
Sipakalaq'bi (Mutual respect)	Respecting others, opinions, and differences.	- Listens attentively when teacher or friends provide explanations; - Values differing opinions in scientific discussions; - Speaks politely to teachers and peers.
Makkeada (Decorum in behavior)	Politeness and etiquette in every action.	- Uses laboratory equipment carefully and orderly; - Raises hand before asking or answering questions; - Maintains calm and order in the classroom; - Speaks politely to teachers, friends, and school residents.
Dipokannyang (Trustworthy/Honesty)	Integrity and consistency between words and deeds.	- Reports experimental data honestly, even if it differs from the hypothesis; - Does not manipulate research data or plagiarize others' work; - Keeps promises, such as submitting assignments on time.

Validation of the affective assessment framework

The framework in Table 5 above was validated through FGD involving teachers, principals, school supervisors, and Mandar customary leaders. The results showed high consensus regarding the relevance of these indicators.

Qualitative data (FGD excerpt): a Mandar customary leader emphasized, "the value of siri' is not just about not stealing, but in the context of learning, children must have siri' to not deceive themselves by committing wrongful acts, and this is very suitable for assessing students' character." A school supervisor added that, "with clear indicators like these, teachers have an objective guide for assessing student attitudes, no longer merely sensing whether a child is good or not."

Teacher-based affective assessment instrument

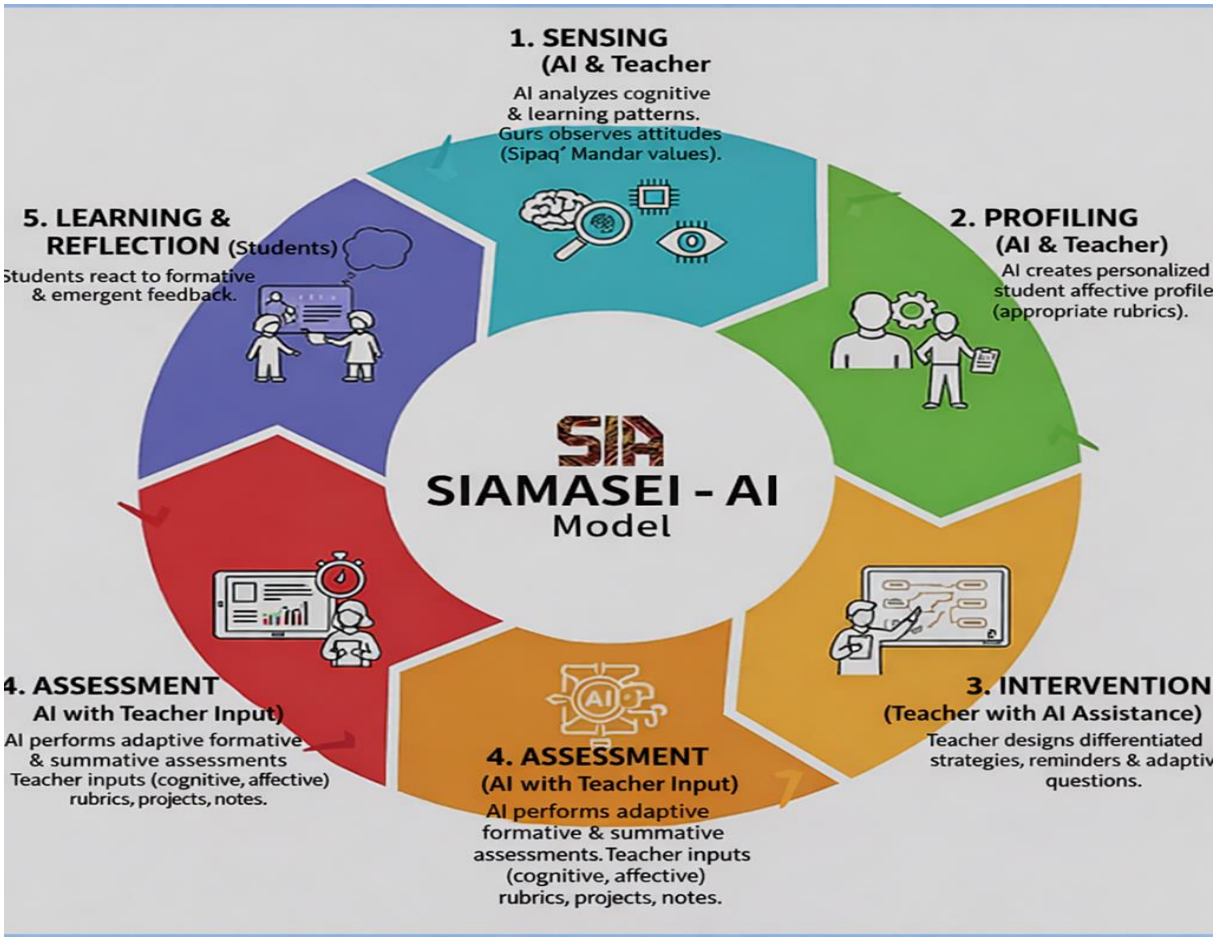
Based on this framework, a teacher-based affective assessment rubric was developed using a 1-4 Likert scale (1= beginning to develop, 2= adequate, 3= good, 4= very good). This rubric is supplemented with an anecdotal record for teachers to note specific examples of student behavior reflecting Sipaq' Mandar values.

Design of the Integrated SIAMASEI-AI Model

Based on a synthesis of all findings, a holistic assessment model named the SIAMASEI-AI Model (Sipaq' Mandar Assessment Integrated with Artificial Learning) was designed. This model illustrates a continuous and synergistic assessment cycle between technology and the teacher's role.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

Figure 1. *SIAMASEI-AI Model*



Source: Processed by researchers (2025) from analysis of research results

Sensing Phase (AI & Teacher): AI analyzes initial data on students' cognitive abilities and learning patterns. The teacher conducts initial observation to map students' attitude tendencies based on Sipaq' Mandar values; Profiling Phase (AI & Teacher): AI creates an individual learning profile (personalized learning profile) covering cognitive-psychomotor strengths and weaknesses. The teacher compiles a student affective profile based on the Sipaq' Mandar rubric;

Intervention Phase (Teacher guided by AI): based on the existing profiles, the teacher designs differentiated learning and assessment strategies. AI recommends materials and questions suitable for each student's cognitive profile and based on clear and appropriate curriculum references and regulations; Assessment Phase (AI & Teacher): AI conducts formative and summative assessment for the cognitive and psychomotor domains adaptively and in real-time. The teacher conducts affective assessment using the Sipaq' Mandar rubric through observation, group project assessment, and anecdotal records; Learning and Reflection Phase (Student & Teacher): students receive a holistic report containing cognitive feedback from AI and affective feedback from the teacher. The teacher and students conduct joint reflection for improvement in the next cycle.

Feasibility and holistic effectiveness of the SIAMASEI-AI model

A two-month trial of the model in science learning (topic: Ecosystems and Environment) provided comprehensive data on the model's feasibility and effectiveness.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

Model feasibility

This feasibility was measured through a teacher perception questionnaire (N=60) using a 1-5 Likert scale.

Table 6. *Results of Model Feasibility Assessment by Teachers*

Assessment Aspect	Average Score	Category
Ease of use (usability) of AI platform	4.2	Highly feasible
Clarity and relevance of Sipaq' Mandar rubric	4.6	Highly feasible
Suitability of the model to the science learning context	4.4	Highly feasible
Model's benefit for the teaching-learning process	4.5	Highly feasible
Overall Average	**4.43**	**Highly feasible**

Qualitative data (teacher interview): one teacher stated, "initially it was a bit complicated, but after implementation, this model is very helpful. I have clear data about student understanding from AI and a strong tool to assess and guide their character. The value of siarondoi is very evident when they work in groups to create ecosystem models."

Model effectiveness

Effectiveness was measured from the improvement in holistic learning achievement (cognitive, psychomotor, and affective) and student perception. Cognitive and psychomotor aspects: as previously outlined, there was a significant increase in post-test scores. ANOVA analysis also showed no significant difference in model effectiveness among the three schools ($F(2,297) = 1.245$, $p = 0.289$), meaning the model is consistently effective across various school contexts in Majene.

Affective aspect: based on teacher rubric assessment data, it shows an increase in the average affective score of students from the first cycle to the second cycle. The average affective score in cycle 1 was 2.8 (category "adequate"), increasing to 3.4 (category "good") in cycle 2. The values of dipokannyang (honesty) and siarondoi (cooperation) showed the highest increase.

Student perception: 88% of students (264 individuals) agreed that the learning outcome report they received (a combination of AI scores and teacher reviews about character) provided a more complete picture of their abilities and self-development. One student said, "now I am not only assessed by numbers, but also by effort, honesty, and how I cooperate. I feel more humanized."

Contribution of AI and cultural values

Multiple linear regression analysis was conducted to examine the contribution of AI application (X1) and the integration of Sipaq' Mandar values (X2) to the effectiveness of holistic assessment (Y). The results showed that both independent variables together provide a significant contribution ($R2 = 0.712$, $p = 0.000$). This means 71.2% of the variation in assessment effectiveness can be explained by the combination of AI use and affective assessment based on Sipaq' Mandar. The Sipaq' Mandar value itself has a higher regression coefficient, indicating that in the context of character formation, the cultural approach has a very strong impact.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

Discussion

This study set out to develop and test an integrated assessment model that bridges the technological imperatives of Society 5.0 with the cultural context of the Mandar community. The following discussion interprets the key findings, elucidates their theoretical and practical significance, acknowledges the study's limitations, and proposes directions for future inquiry.

Interpretation of key findings

The findings collectively affirm that a synergistic integration of AI and local wisdom-based assessment is not only feasible but also effective in enhancing holistic learning evaluation. This addresses the core integrative gap identified in the literature, where technological and cultural assessment approaches have historically existed in parallel (Awaru et al., 2024; Dubey et al., 2025).

First, the significant improvement in cognitive-psychomotor scores (Δ=11.7,$p<.001$,d=1.12), corroborated by qualitative feedback on AI's real-time feedback and adaptive questioning, confirms the role of AI as an effective augmenting tool for teachers (Hendri et al., 2025). It automates routine measurement, allowing teachers to reallocate their energy towards more complex pedagogical tasks, thus aligning with the vision of AI as a partner rather than a replacement (El Karafli, 2025).

Second, and more crucially, the successful operationalization of Sipaq' Mandar values into observable behavioral indicators (Table 5) provides a concrete bridge between abstract cultural philosophy and everyday classroom assessment. This process directly tackles the operationalization gap. The high consensus from FGDs with cultural leaders and educators validates that values like siri' (shame/honor) and siarondoi (cooperation) are not antiquated concepts but are highly relevant to modern educational goals like academic integrity and collaborative problem-solving. This empowers teachers to act as cultural translators and moral agents (Fepriyanti & Roqib, 2024), assessing the "human dimension" of learning that AI cannot reach.

Third, the SIAMASEI-AI Model (Figure 1) crystallizes the answer to the need for a structured integration framework. Its cyclical, five-phase design creates a symbiotic ecosystem: AI handles data-intensive profiling and cognitive assessment, while teachers lead the interpretive, value-based affective assessment and intervention. This mutualistic relationship ensures assessment is a continuous, formative process embedded in learning, not a disconnected summative event (Pandya, 2024).

Finally, the model's high feasibility (M=4.43/5) and its substantial contribution to holistic assessment effectiveness (R^2 =.712) empirically validate its utility in a real-world context. The stronger regression coefficient for Sipaq' Mandar values compared to the AI component underscores a pivotal insight: in the Society 5.0 era, the cultural foundation is not diluted by technology but is instead the key to humanizing and grounding technological application. Students' reported feeling "more humanized" when assessed holistically is a powerful testament to this balance.

Theoretical implications: towards a cultural-TPACK framework

This study makes a significant theoretical contribution by proposing and empirically validating an extension of the classic TPACK framework (Mishra & Koehler, 2006). Cultural-TPACK. While TPACK emphasizes the intersection of Technological, Pedagogical, and Content Knowledge, our findings reveal that in culturally rich and technologically evolving settings like Indonesia, Local Wisdom Knowledge (LWK) is a critical fourth domain.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

The development and successful implementation of the SIAMASEI-AI Model demonstrate that effective teaching and assessment in the Society 5.0 require teachers to navigate not just TK, PK, and CK, but also LWK. Teachers must understand how to use an AI platform (TK) to assess science concepts (CK) through inquiry methods (PK) while simultaneously interpreting student behaviors through the lens of sipakalaq'bi (mutual respect) or dipokannyang (honesty) (LWK). The Cultural-TPACK framework, illustrated in Figure 3 (to be included), posits that the most effective and contextually appropriate practice lies at the dynamic intersection of all four knowledge domains. This framework provides a new lens for teacher professional development and technology integration policies, moving beyond a technocentric view to a socio-techno-cultural perspective.

Practical implications for teachers and policy

The operational outputs of this research offer direct practical utility:

1. **For Teachers and Schools:** The validated Sipaq' Mandar affective assessment rubric (Table 5) is a ready-to-use instrument. Schools can adopt the SIAMASEI-AI cycle incrementally, starting with subjects like science. Continuous professional development should focus on building teachers' Cultural-TPACK, combining AI tool literacy with training in ethnographic observation and anecdotal note-taking for character assessment.
2. **For Educational Policymakers:** This model offers a blueprint for implementing the Merdeka Curriculum, which emphasizes character strengthening and local content. District and national education offices can integrate the principles of the SIAMASEI-AI Model into assessment guidelines, promoting "glocal" assessment standards—globally competent yet locally rooted. Funding should be directed towards building technological infrastructure in tandem with programs that document and operationalize local wisdom.
3. **For Curriculum Designers:** Learning modules, especially in science and social studies, can be designed with embedded tasks that naturally elicit behaviors tied to local values (e.g., group projects that require siarondoi), making affective assessment more authentic and less burdensome.

Limitations of the study

While promising, this study has limitations that must be acknowledged. First, the research was conducted in a specific geographical and cultural context (Mandar community in Majene). The direct transferability of the Sipaq' Mandar rubric to other ethnic groups in Indonesia requires adaptation and re-validation. Second, the duration of the model trial was two months within a single topic (Ecosystems). The long-term sustainability of the effects on character development and the model's efficacy across different school subjects remain to be tested. Third, the study relied on available school technology infrastructure. The scalability of the model to schools with limited digital resources presents a significant practical challenge. Finally, the affective assessment, though rubric-based, still involves teacher subjectivity. While the rubric standardizes the criteria, inter-rater reliability over a longer period and across many teachers needs further monitoring.

Suggestions for future research

Based on these findings and limitations, future research should:

1. Conduct longitudinal studies to examine the sustained impact of the SIAMASEI-AI Model on students' character development, academic resilience, and overall school climate over one or more academic years.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

2. Explore adaptation and replication in diverse cultural contexts across Indonesia (e.g., applying the framework with Papua, Javanese, or Batak local wisdom) to develop a more generalized yet flexible model for culturally responsive assessment.
3. Investigate the specific development of teacher competency in Cultural-TPACK. Research could design and evaluate training modules aimed at enhancing teachers' synergistic mastery of technology, pedagogy, content, and local wisdom knowledge.
4. Develop and test lightweight, offline-capable AI tools or simplified analog-digital hybrid systems to make the technological component of such models accessible to schools in remote or under-resourced areas, ensuring equitable implementation.

Conclusion and recommendations

This study concludes that the integration of Artificial Intelligence with culturally-grounded affective assessment offers a viable, balanced, and humanized evaluation framework for the Society 5.0 era. The SIAMASEI-AI model effectively synergizes technological objectivity with the irreplaceable role of teachers as moral agents and cultural translators. The key takeaways from this research are summarized below.

Key Findings with specific evidence

The research yielded robust evidence supporting the model's effectiveness. First, the integration of AI for cognitive-psychomotor assessment led to a statistically significant improvement in student scores, with a mean increase of 11.7 points (from 68.5 to 80.2, $p < .001$, Cohen's $d = 1.12$). Second, the five core Sipaq' Mandar values were successfully operationalized into a valid and reliable affective assessment rubric, which was deemed highly feasible by teachers ($M = 4.43/5$). Third, multiple regression analysis demonstrated that the combined use of AI and Sipaq' Mandar values contributed significantly (71.2%, $R^2 = 0.712$) to the effectiveness of holistic assessment, with the cultural component showing a stronger influence.

Theoretical contributions

The primary theoretical contribution of this study is the proposal and empirical validation of the Cultural-TPACK framework. By adding Local Wisdom Knowledge (LWK) as a fourth domain to the existing TPACK model, this research provides a more nuanced theoretical lens for understanding and developing teacher competencies in technologically advanced, culturally diverse classrooms. This framework emphasizes that effective education in the 21st century is not merely about integrating technology, but about integrating it in a way that is pedagogically sound and culturally responsive.

Practical implications

This research provides direct, actionable outputs for educational stakeholders. For teachers and schools, the SIAMASEI-AI model and its accompanying Sipaq' Mandar rubric serve as a ready-to-use blueprint for conducting holistic assessments that are both efficient and meaningful. For educational policymakers, the findings offer a strategic model for implementing the character-building goals of the Merdeka Curriculum, advocating for "glocal" assessment standards that are globally competitive yet deeply rooted in local wisdom.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

Limitations of the study

Despite its contributions, this study has limitations. The research was confined to the specific cultural context of the Mandar community in Majene, which may limit the generalizability of the affective rubric to other ethnic groups. The model's trial period was also relatively short (two months), so its long-term effects on character development remain untested. Furthermore, the model's reliance on existing technological infrastructure could pose a challenge for scaling up in under-resourced schools, and the affective assessment, while rubric-based, still involves an element of teacher subjectivity.

Future research directions

Building on these findings, future research should pursue several avenues. Longitudinal studies are needed to assess the sustained impact of the SIAMASEI-AI model on students' character and school climate. Research should also explore the adaptation and implementation of the Cultural-TPACK framework in diverse cultural contexts across Indonesia. Furthermore, investigations into designing and evaluating teacher professional development programs focused on building Cultural-TPACK competencies, as well as developing lightweight, offline-compatible AI tools, are crucial for ensuring the equitable and sustainable adoption of such integrated models.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Agrawal, A., Kumar, S., & Singh, P. (2024). Smart assessment systems: A paradigm shift in education. Journal of Educational Technology & Society, 27(3), 45–58.

Awaru, M., Rasyid, M., & Ahmad, A. (2024). Sibaliparriq values as the foundation of character education in the Mandar community. Jurnal Pendidikan Kewarganegaraan, 10(1), 78-92.

Black, P., & Wiliam, D. (2009). Assessment for learning: Why it matters and how it can make a difference. Routledge.

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. David McKay Co.

Branch, R. M. (2009). Instructional design: The ADDIE approach. Springer Science & Business Media.

Burhanuddin, B., Alimin, M., & Rasyid, R. (2023). The philosophical meaning of Pappasang and Sipaq' Mandar in building social resilience. Jurnal Studi Budaya, 15(2), 201-218.

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). Sage publications.

Data of Education Office. (2024). Laporan implementasi teknologi pendidikan pasca-gempa dan pandemi di Provinsi Sulawesi Barat Tahun 2024 [Report on the implementation of educational technology post-earthquake and pandemic in West Sulawesi Province for the year 2024]. Education Office of West Sulawesi Province.

Dubey, A., Singh, P., & Kumar, R. (2025). Adaptive learning platforms powered by AI: A comprehensive analysis of student performance metrics. Computers & Education, 198, 104-115.

ISSN |2355-3669| E-ISSN |2503-2518|
Volume 13 | Number 1 | June 2026|

El Karafli, S. (2025). The human element in the age of AI: Safeguarding affective domains in education. In A. K. Singh & L. Chen (Eds.), Bridging humanities and technology in modern pedagogy (pp. 112-128). Academic Press.

Fepriyanti, Y., & Roqib, M. (2024). The irreplaceable role of teachers as moral agents in technology-infused classrooms. Jurnal Pendidikan dan Kebudayaan, 30(1), 15-28.

Gay, G. (2018). Culturally responsive teaching: Theory, research, and practice (3rd ed.). Teachers College Press.

Hendri, H., Susanto, A., & Wijaya, T. (2025). AI-driven feedback in science education: An empirical study on its impact on conceptual change. International Journal of Science Education, 47(4), 567-585.

Kumar, V. (2025). Automated scoring systems for performance-based assessments: A review of accuracy and bias. Educational Technology Research and Development, 73(1), 1-21.

Matheis, C., & John, J. (2024). Beyond the cognitive: Assessing 21st-century competencies in a digital world. Journal of Learning Analytics, 11(2), 89-105.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017–1054.

Pandya, J. (2024). Society 5.0 and the cognitive partnership: Redefining the role of technology in learning design. TechTrends, 68(1), 12-23.

Prestoza, J. R., Reyes, M. L., & Tan, F. T. (2025). Data-driven instruction: Leveraging learning analytics for pedagogical decision-making. British Journal of Educational Technology, 56(3), 612-629.

Research Data. (2025). Preliminary data on assessment needs in Majene Regency [Unpublished raw data]. Universitas Sulawesi Barat.

Sánchez Rogel, M. J., Pérez-García, M., & Fernández, L. (2024). Teachers as cultural translators: A qualitative study on affective assessment in diverse classrooms. Teaching and Teacher Education, 140, 104-116.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Yunus, A. (2025). Preliminary data on assessment needs in Majene Regency [Unpublished raw data]. Universitas Sulawesi Barat.