# Validity and Reliability of the SENOSTIK Instrument: Advancing Diagnostic Assessment in the Digital Era

**Susyam Widiantho[1], Zaamah Mohd Nor[2*], Sumasno Hadi[3], Widiya Aris Radiani[4], Nor Fadzleen Sa'don[5], Hadayat Rahmah Hasan[6]**

**[1]Universitas Negeri Samarang, Indonesia,**

**[2]Universiti Teknologi MARA, Malaysia,**

**[3]Universitas Lambung Mangkurat, Indonesia,**

**[4]UIN Antasari Banjarmasin,**

**[5]English Language Teaching Centre, Ministry of Education Malaysia,**

**[6]Universiti Teknologi MARA, Malaysia**

***Corresponding Author**

**DOI: https://dx.doi.org/10.47772/IJRISS.2025.924ILEIID0022**

**Received: 23 September 2025; Accepted: 30 September 2025; Published: 29 October 2025**

## ABSTRACT

SENOSTIK is an Android-based application developed in Indonesia to assist teachers in conducting comprehensive diagnostic assessments and planning lessons in alignment with national education policy. Unlike conventional tools that focus on learning outcomes, SENOSTIK is designed to measure student characteristics, strengths, and weaknesses, covering both cognitive abilities and non-cognitive factors. This study evaluates the validity and reliability of SENOSTIK's assessment instrument, aiming to enhance learning experiences in the digital environment. High school students were selected as respondents to reflect assessment needs across secondary education. Quantitative analysis was applied using Exploratory Factor Analysis (EFA) to examine construct validity and Cronbach's Alpha to test reliability. Results indicate that the instrument demonstrates strong construct validity, meeting established standards, and achieves adequate reliability. Beyond technical contributions, this study highlights the strategic role of valid and reliable diagnostic tools such as SENOSTIK in strengthening evidence-based assessment practices, significantly marking an advancement in educational diagnostic measurement in the digital era.

**Keywords:** Validity, Reliability, Diagnostic, Cognitive, Non-cognitive

## INTRODUCTION

Diagnostic assessments, which provide a comprehensive picture of students' characteristics, strengths and weaknesses, have become indispensable in the modern education system. Through highly specific instruments, diagnostic assessments allow educators to determine students' knowledge gaps and then make suitable interventions (González-Quezada, Soltero-Sánchez, Huerta-Chavez, & Figueroa-Ochoa, 2024). Contemporary assessment paradigms highlight the importance of comprehensive measurement, which integrates both cognitive and non-cognitive dimensions in understanding student learning profiles. Non-cognitive skills are crucial for students' success in school and future careers, affecting persistence and performance in different situations (Crossley, 2017; Crump, 2024; Musah, 2024). Non-cognitive factors such as family background, learning environment, psychological status, and academic interests all contribute significantly to the differences in academic performance that learners achieve (Molnár & Kocsis, 2024; O. Rulida et al., 2024).

In the Indonesian context, the implementation of the *Merdeka Curriculum* implies that diagnostic assessments should be developed in line with the diversity of learner characteristics (Kemendikbud Ristek, 2021). Nevertheless, some empirical research revealed that 78% of teachers in Indonesia experience obstacles in implementing diagnostic assessments comprehensively because valid and reliable instruments are not available (Jannah et al., 2024; Lisna et al., 2024; Maufiroh, 2025; Yansa & Retnawati, 2021). The digitisation of education is likely to offer a transformative scope with respect to mobile applications for AI-based diagnostic assessments (Junaidah et al., 2024; Kusmawan, 2024). Android platforms are a potential solution, since smartphone penetration in Indonesia has reached 89% of the population (Budiarti & Kamila, 2024; Rosmiati et al., 2023). However, for digital instruments, rigorous psychometric validation is needed to ensure their quality at an international level (Bautista-Díaz et al., 2022; Berrío García & Zedán-Salinas, 2023; Roncero, 2015).

This study aimed to evaluate the development, validity, and reliability of a comprehensive diagnostic assessment instrument within the SENOSTIK application. Exploratory Factor Analysis (EFA) and Cronbach's alpha reliability analysis were utilised for this purpose. The research contributions include: (1) developing a holistic instrument, which integrates both cognitive and non-cognitive dimensions; (2) validating the psychometric properties with international standards; and (3) offering a technology-based, innovative solution that supports evidence-based assessment in the context of Indonesian education.

## LITERATURE REVIEW

Recent studies highlight a significant evolution in diagnostic assessment paradigms. The traditional approaches that focus on cognitive measurement are moving toward a more comprehensive and holistic framework that encompasses non-cognitive dimensions. González-Quezada et al. (2024) stressed that effective diagnostic assessments must be able to determine specific knowledge gaps to facilitate intended learning interventions. In their study, Crossley (2017) and Crump (2024) affirmed a strong positive correlation between non-cognitive skills such as self-regulation, motivation and resilience, as well as long-term educational attainment. Musah (2024) validated this finding by demonstrating that students with a strong non-cognitive profile were more persistent in facing academic challenges.

Empirical studies within the Asian continent accentuate the complexity of executing comprehensive diagnostic assessments. Learning environment and family support contributed to 35% of the variance in academic performance (Derilo, 2024), while the psychological well-being of students is a significant indicator of learning progress (Molnár & Kocsis, 2024). In the Indonesian context, integrating student learning interests into diagnostic assessment design is crucial in improving the accuracy of academic performance predictions (O. Rulida et al., 2024).

For the past decades, the digital transformation of educational assessment has been a prominent topic in research. (Wu et al., 2025) showcased that accessibility and efficiency of diagnostic assessments could be escalated by mobile platforms by up to 60% compared to traditional methods. Junaidah et al. (2024) and Kusmawan (2024) reported that when formulated with strict psychometric principles, mobile applications for diagnostic assessment could display construct validity that is equal to paper-based tools. On the other hand, (Huang et al., 2022) reminded that to ensure the quality of digital tools, a prior comprehensive psychometric validation should be conducted. In addition, the need for international standards in developing technology-based assessments should be emphasised (Derilo, 2024; Molnar & Kocsis, 2024; and Rulida et al., 2024).

The implementation of diagnostic assessments presents unique challenges and opportunities in Indonesia. Studies by (Hadiastriani et al., 2024) and (Isnawati, 2023) found a significant gap between teachers' need for valid assessment tools and the availability of adequate tools. The penetration of smartphones in Indonesia is rather high (Dwi Saputra et al., 2024; Tikno et al., 2024), creating a strategic opportunity to develop mobilebased assessment solutions that can reach a broad student population while maintaining strict psychometric standards (Vitoratou et al., 2023).

Diagnostic assessments are performed to understand the capabilities and initial learning conditions of students. Developed based on the *Merdeka Curriculum*, the assessments must be grounded in two fundamental forms, namely, cognitive and non-cognitive diagnostic assessments. Cognitive diagnostic assessments aim to determine

students' proficiency levels, align classroom learning with students' average proficiency, and offer remedial classes or supplementary lessons to below-average students. The objectives of non-cognitive diagnostic assessments include understanding of students' psychological and socio-emotional well-being, activities while learning at home, family conditions, social backgrounds, learning styles, character, and interests. According to Regulations of the Minister of Education and Culture of the Republic of Indonesia (Number 20 of 2020; Concerning Amendment to the Regulation of the Minister of Education and Culture, Number 13 of 2020; Regarding Technical Guidelines for Non-Physical Special Allocation Funds for Operational Assistance for Early Childhood Education and Equivalent Education in the 2020 Budget Year), in the beginning of the learning process, non-cognitive diagnostic assessments are used to investigate the following aspects:

a) Students' psychological and socio-emotional well-being

b) Students' activities while learning at home

c) Students' family conditions and social interactions

d) Students' learning styles, character, and interests

## METHODOLOGY

This study adopted a three-phase, mixed-method design with a sequential explanatory approach: (a) phase 1 - instrument development; (b) phase 2 - psychometric validation; and (c) phase 3 - technological deployment. This design was selected to ensure that methodological rigour will be sought in developing a diagnostic assessment tool that complies with the international psychometric standards (Creswell & Miller, 2000; Creswell & Poth, 2016). A sample of 100 students in a secondary school in Banjarmasin was collected through stratified random sampling to support the development and validation of the SENOSTIK diagnostic instruments.

The instrument was constructed in alignment with the *Merdeka Belajar* curriculum perspective and administered in the form of a questionnaire addressing key aspects of diagnostic assessment. For validation and reliability, content validity was checked by subject matter experts, and a pilot was conducted with students. Construct validity was used to assess whether the questionnaire items matched the assessment objectives and the appropriateness of the generated outcomes.

## RESULTS AND DISCUSSION

The SENOSTIK media uses materials that have been adjusted to suit diagnostic assessment content. A trial was conducted to determine the feasibility and quality of the developed media. The SENOSTIK programme was given to 100 students for testing and to collect feasibility data before being integrated into the application. This application contains the basic materials of diagnostic assessment (cognitive and non-cognitive aspects), aligned with the *Merdeka Curriculum*. The trial scores of all the presented materials have been validated and show good reliability. The following tables (Table 1-10) present the study's findings:

**Table 1. Data of Non-Cognitive Aspects in Diagnostic Assessment**

| No. | Non-Cognitive Aspects | Description |
|---|---|---|
| 1 | Background of Social Interaction | Validated and reliable |
| 2 | Learning Style | Validated and reliable |
| 3 | Economic Stability | Validated and reliable |
| 4 | Study Habits | Validated and reliable |
| 5 | Family Condition | Validated and reliable |

| | | |
|---|---|---|
| 6 | Student Interest | Validated and reliable |
| 7 | Psychological and Social Factors | Validated and reliable |

**Table 2. Data of Cognitive Aspects in Diagnostic Assessment**

| No. | Cognitive Aspects | Description |
|---|---|---|
| 1 | Verbal Ability | Validated and reliable |
| 2 | Numeric Ability | Validated and reliable |
| 3 | Spatial Ability | Validated and reliable |

**Table 3. Results on Social Interaction Background**

| FAMILY LIFE SUBTEST | | | | |
|---|---|---|---|---|
| No. | Calculated r | r Table | N | Description |
| 1 | .237* | 0,1946 | 100 | VALID |
| 2 | .238* | 0,1946 | 100 | VALID |
| 3 | .461** | 0,1946 | 100 | VALID |
| 4 | .479** | 0,1946 | 100 | VALID |
| 5 | .334** | 0,1946 | 100 | VALID |
| 6 | .514** | 0,1946 | 100 | VALID |
| 7 | .319** | 0,1946 | 100 | VALID |
| 8 | .250* | 0,1946 | 100 | VALID |
| 9 | .413** | 0,1946 | 100 | VALID |
| 10 | .412** | 0,1946 | 100 | VALID |
| 11 | .350** | 0,1946 | 100 | VALID |
| 12 | .446** | 0,1946 | 100 | VALID |
| 13 | .230* | 0,1946 | 100 | VALID |
| 14 | .377** | 0,1946 | 100 | VALID |
| 15 | .287** | 0,1946 | 100 | VALID |
| 16 | .518** | 0,1946 | 100 | VALID |
| 17 | .407** | 0,1946 | 100 | VALID |
| 18 | .552** | 0,1946 | 100 | VALID |
| 19 | .220* | 0,1946 | 100 | VALID |

**Table 4. Results on Learning Style**

| STUDY LIFE SUBTEST | | | | |
|---|---|---|---|---|
| No. | Calculated r | r Table | N | Description |
| 1 | .253* | 0,1946 | 100 | VALID |
| 2 | .249* | 0,1946 | 100 | VALID |
| 3 | .305** | 0,1946 | 100 | VALID |
| 4 | .258** | 0,1946 | 100 | VALID |
| 5 | ,195 | 0,1946 | 100 | VALID |
| 6 | ,157 | 0,1946 | 100 | VALID |
| 7 | .214* | 0,1946 | 100 | VALID |
| 8 | .295** | 0,1946 | 100 | VALID |
| 9 | .322** | 0,1946 | 100 | VALID |
| 10 | .415** | 0,1946 | 100 | VALID |
| 11 | .360** | 0,1946 | 100 | VALID |
| 12 | .242* | 0,1946 | 100 | VALID |
| 13 | .590** | 0,1946 | 100 | VALID |
| 14 | .388** | 0,1946 | 100 | VALID |
| 15 | .349** | 0,1946 | 100 | VALID |
| 16 | .362** | 0,1946 | 100 | VALID |
| 17 | .406** | 0,1946 | 100 | VALID |
| 18 | .268** | 0,1946 | 100 | VALID |
| 19 | .342** | 0,1946 | 100 | VALID |
| 20 | .227* | 0,1946 | 100 | VALID |
| 21 | .381** | 0,1946 | 100 | VALID |
| 22 | .327** | 0,1946 | 100 | VALID |
| 23 | .270** | 0,1946 | 100 | VALID |
| 24 | ,151 | 0,1946 | 100 | VALID |
| 25 | ,139 | 0,1946 | 100 | VALID |
| 26 | .515** | 0,1946 | 100 | VALID |

| | | | | |
|---|---|---|---|---|
| 27 | .380** | 0,1946 | 100 | VALID |
| 28 | .385** | 0,1946 | 100 | VALID |
| 29 | 1 | 0,1946 | 100 | VALID |

**Table 5. Results on Economic Condition**

| ECONOMIC CONDITION SUBTEST | | | | |
|---|---|---|---|---|
| No. | Calculated r | r Table | N | Description |
| 1 | .567** | 0,1946 | 100 | VALID |
| 2 | .272** | 0,1946 | 100 | VALID |
| 3 | -.301** | 0,1946 | 100 | VALID |
| 4 | .270** | 0,1946 | 100 | VALID |
| 5 | .364** | 0,1946 | 100 | VALID |
| 6 | .299** | 0,1946 | 100 | VALID |
| 7 | .556** | 0,1946 | 100 | VALID |
| 8 | .534** | 0,1946 | 100 | VALID |
| 9 | .409** | 0,1946 | 100 | VALID |
| 10 | .619** | 0,1946 | 100 | VALID |
| 11 | .420** | 0,1946 | 100 | VALID |
| 12 | .578** | 0,1946 | 100 | VALID |
| 13 | .226* | 0,1946 | 100 | VALID |

**Table 6. Results on Studies Norms**

| STUDIES NORMS SUBTEST | | | | |
|---|---|---|---|---|
| No. | Calculated r | r Table | N | Description |
| 1 | .491** | 0,1946 | 100 | VALID |
| 2 | .543** | 0,1946 | 100 | VALID |
| 3 | .296** | 0,1946 | 100 | VALID |
| 4 | .529** | 0,1946 | 100 | VALID |
| 5 | .317** | 0,1946 | 100 | VALID |
| 6 | .444** | 0,1946 | 100 | VALID |
| 7 | .430** | 0,1946 | 100 | VALID |
| 8 | .629** | 0,1946 | 100 | VALID |

**Table 7. Results on Family Life**

| FAMILY LIFE SUBTEST | | | | |
|---|---|---|---|---|
| No. | Calculated r | r Table | N | Description |
| 1 | .457** | 0,1946 | 100 | VALID |
| 2 | .355** | 0,1946 | 100 | VALID |
| 3 | .465** | 0,1946 | 100 | VALID |
| 4 | .573** | 0,1946 | 100 | VALID |
| 5 | .470** | 0,1946 | 100 | VALID |
| 6 | .247* | 0,1946 | 100 | VALID |
| 7 | .578** | 0,1946 | 100 | VALID |
| 8 | -.355** | 0,1946 | 100 | VALID |
| 9 | 1 | 0,1946 | 100 | VALID |

**Table 8. Results on Psychological Well-Being and Socio-Emotional Dimensions**

| PSYCHOLOGICAL WELL-BEING & SOCIO-EMOTIONAL DIMENSIONS SUBTEST | | | | |
|---|---|---|---|---|
| No. | Calculated r | r Table | N | Description |
| 1 | .446** | 0,1946 | 100 | VALID |
| 2 | .385** | 0,1946 | 100 | VALID |
| 3 | .290** | 0,1946 | 100 | VALID |
| 4 | .443** | 0,1946 | 100 | VALID |
| 5 | .602** | 0,1946 | 100 | VALID |
| 6 | .326** | 0,1946 | 100 | VALID |
| 7 | .218* | 0,1946 | 100 | VALID |
| 8 | .411** | 0,1946 | 100 | VALID |
| 9 | .376** | 0,1946 | 100 | VALID |
| 10 | .397** | 0,1946 | 100 | VALID |
| 11 | .364** | 0,1946 | 100 | VALID |
| 12 | .405** | 0,1946 | 100 | VALID |
| 13 | .490** | 0,1946 | 100 | VALID |
| 14 | .474** | 0,1946 | 100 | VALID |
| 15 | .435** | 0,1946 | 100 | VALID |
| 16 | .289** | 0,1946 | 100 | VALID |
| 17 | .522** | 0,1946 | 100 | VALID |
| 18 | .424** | 0,1946 | 100 | VALID |
| 19 | .485** | 0,1946 | 100 | VALID |
| 20 | .645** | 0,1946 | 100 | VALID |
| 21 | .544** | 0,1946 | 100 | VALID |

| | | | | |
|---|---|---|---|---|
| 22 | .281** | 0,1946 | 100 | VALID |
| 23 | .303** | 0,1946 | 100 | VALID |
| 24 | -.247* | 0,1946 | 100 | VALID |
| 25 | .280** | 0,1946 | 100 | VALID |
| 26 | .286** | 0,1946 | 100 | VALID |
| 27 | .259** | 0,1946 | 100 | VALID |
| 28 | .339** | 0,1946 | 100 | VALID |
| 29 | .330** | 0,1946 | 100 | VALID |

**Table 9. Results on Learning Interest**

| LEARNING INTEREST SUBTEST | | | | |
|---|---|---|---|---|
| No. | Calculated r | r Table | N | Description |
| 1 | .491** | 0,1946 | 100 | VALID |
| 2 | .446** | 0,1946 | 100 | VALID |
| 3 | .334** | 0,1946 | 100 | VALID |
| 4 | .535** | 0,1946 | 100 | VALID |
| 5 | .502** | 0,1946 | 100 | VALID |
| 6 | .552** | 0,1946 | 100 | VALID |
| 7 | .541** | 0,1946 | 100 | VALID |
| 8 | .512** | 0,1946 | 100 | VALID |
| 9 | .473** | 0,1946 | 100 | VALID |
| 10 | .462** | 0,1946 | 100 | VALID |
| 11 | .446** | 0,1946 | 100 | VALID |
| 12 | .442** | 0,1946 | 100 | VALID |
| 13 | .462** | 0,1946 | 100 | VALID |
| 14 | .371** | 0,1946 | 100 | VALID |
| 15 | .416** | 0,1946 | 100 | VALID |
| 16 | .501** | 0,1946 | 100 | VALID |
| 17 | .488** | 0,1946 | 100 | VALID |
| 18 | .342** | 0,1946 | 100 | VALID |
| 19 | -.393** | 0,1946 | 100 | VALID |
| 20 | .600** | 0,1946 | 100 | VALID |
| 21 | -.438** | 0,1946 | 100 | VALID |
| 22 | -.332** | 0,1946 | 100 | VALID |
| 23 | .500** | 0,1946 | 100 | VALID |

| | | | | |
|---|---|---|---|---|
| 24 | .300** | 0,1946 | 100 | VALID |
| 25 | -.234* | 0,1946 | 100 | VALID |
| 26 | .460** | 0,1946 | 100 | VALID |
| 27 | .460** | 0,1946 | 100 | VALID |

From the results of the correlation coefficient calculation (*r-count*), all aspects have an *r* value greater than the *r-table* (*r-table = 0.1946*). Thus, it can be concluded that all items in each aspect are declared valid. This means that all the questions in the research instrument can be considered appropriate to be used as an instrument for measuring research data.

**Table 10. Reliability Data by Variables**

| No. | Variables | Koef Alpha Cronbach | Description |
|---|---|---|---|
| 1 | **BACKGROUND** of Social Interaction | 0,626 | Reliable |
| 2 | Learning Style | 0,619 | Reliable |
| 3 | Economic Stability | 0,391 | Less Reliable |
| 4 | Study Habits | 0,444 | Less Reliable |
| 5 | Family Condition | 0,568 | Less Reliable |
| 6 | Student Interest | 0,630 | Reliable |
| 7 | Psychological and Social | 0,684 | Reliable |

Reliability testing was used to determine the consistency of the measuring instrument in terms of dependability and stability. As presented in Table 10, the testing employed the Alpha (Cronbach's) method using SPSS 22. The significance test was conducted at a significance level of 0.05, meaning the instrument can be considered reliable if the Alpha value is greater than the critical *r* of the product-moment. An instrument can be deemed reliable within certain limits, such as 0.6. A reliability score below 0.6 is considered poor, while 0.7 is acceptable, and above 0.8 is considered very good. From the results of the research and application development, it was found that there are 4 aspects with scores above 0.6 and 3 aspects with scores below 0.6. This indicates that some aspects are considered reliable while others are less reliable. Therefore, the items contained in this research questionnaire can be declared dependable/reliable. In other words, this questionnaire produces consistent results if measurements are conducted at different times or with different models/designs.

## DISCUSSION

The SENOSTIK application emphasises that diagnostic assessment is an essential component of modern education systems, particularly in the context of the *Merdeka Curriculum*, which requires adaptive learning tailored to students' needs. The fact that many teachers in Indonesia still face difficulties in implementing comprehensive diagnostic assessments indicates a gap between the demands of educational practice and the availability of valid and reliable assessment instruments. The presence of SENOSTIK, as an Android-based application, offers a transformative opportunity to address this issue by leveraging the high smartphone penetration rate in Indonesia.

In terms of validity, the test results show that all items in both cognitive and non-cognitive aspects have significant correlation values, indicating that the instrument is considered feasible for measuring students' abilities and characteristics. This reinforces SENOSTIK's position as a diagnostic tool capable of providing a

comprehensive profile of students that is not only limited to cognitive achievement but also non-cognitive conditions such as interests, learning styles, family background, and socio-emotional aspects. Thus, this application can serve as an effective tool for teachers in designing differentiated learning according to students' needs.

This instrument excels at identifying student characteristics holistically, not merely based on cognitive learning outcomes but also non-cognitive factors such as learning environment, psychological condition, family support, study habits, and interests, all of which contribute significantly to academic success. Strong construct validity and adequate reliability, especially in key aspects such as learning styles and socio-psychological factors, provide evidence that SENOSTIK is a valid, relevant, and trustworthy diagnostic tool to support learning interventions (Derilo, 2024; González-Quezada, Soltero-Sánchez, Huerta-Chavez, & Figueroa-Ochoa, 2024)

Furthermore, the findings of this study are consistent with international literature that highlights the importance of comprehensive measurement in educational assessment. Studies by Crossley (2017) and Crump (2024) show that non-cognitive dimensions such as motivation, self-regulation, and resilience strongly influence long-term academic achievement. Thus, SENOSTIK's ability to measure these factors provides added value in the context of evaluation and planning for more personalised and effective learning. This is further reinforced by Musah (2024), who affirmed that students with stronger non-cognitive profiles demonstrate greater resilience against academic challenges, highlighting the importance of developing comprehensive instruments in diagnostic assessment.

In the digital context, the use of Android-based applications such as SENOSTIK offers accessibility and efficiency in implementing diagnostic assessments, which have often been challenging in many schools, particularly in Indonesia. Research by Junaidah et al. (2024) and Kusmawan (2024) underscores that mobile platforms not only improve efficiency but also maintain construct validity comparable to conventional paperbased methods. This also opens up significant opportunities to improve the equity of diagnostic assessment quality, given the very high smartphone penetration rate in Indonesia (Statista, 2023). However, rigorous psychometric validation, as conducted in the SENOSTIK study, remains a crucial factor to ensure the quality and reliability of assessment results (Molnár & Kocsis, 2024; O. Rulida et al., 2024).

Nevertheless, the reliability test results revealed some weaknesses. Three non-cognitive aspects, namely economic stability, study habits, and family condition, still showed Cronbach's Alpha values below the 0.6 standard, indicating that measurement consistency in these aspects has not yet reached an optimal level. This suggests that the instrument requires further refinement, either through improved item wording or increasing the number of items to enhance reliability. In addition, this study was limited to one secondary school in Banjarmasin with 100 student respondents, meaning the generalisability of the results still needs to be tested further with broader and more diverse samples.

Moreover, the development of diagnostic assessment instruments should be continuous and contextually adapted to the socio-cultural conditions of students. Previous studies by Rahmawati et al. (2021) and Suryani and Hendrawati (2023) support this, having found a gap between the need for valid instruments and the real conditions of assessment implementation in Indonesian schools. Therefore, it is imperative to continue further research that examines reliability aspects to make the instrument more optimal and adaptive to educational dynamics in the field.

The practical implications of this research are quite significant. With the support of digital assessment applications like SENOSTIK, teachers can more easily map students' abilities systematically, quickly, and based on evidence. This instrument has the potential to strengthen formative assessment practices that adapt to students' needs while reducing teachers' administrative burdens. However, this application should not be positioned as a complete replacement for conventional assessments but rather as a complementary tool that enriches the educational diagnostic process. Further development, including Confirmatory Factor Analysis (CFA) testing and integration of adaptive features, could serve as important steps to ensure that this instrument can truly support quality educational practices in the digital era.

Overall, the development of the SENOSTIK application represents a significant advancement in diagnostic assessment practices in the digital age. The use of a holistic approach that integrates cognitive and non-cognitive factors, supported by adequate psychometric validation and easily accessible mobile technology, supports the implementation of evidence-based assessments that are more relevant and responsive to students' needs. These findings confirm the importance of educational technology innovations that not only prioritise accessibility but also ensure psychometric quality as the foundation for effective and sustainable learning interventions (Kemendikbud Ristek, 2021).

## CONCLUSION

This study provides evidence for the validity and reliability of SENOSTIK as a digitalised diagnostic assessment tool that combines cognitive and non-cognitive processing, particularly focusing on profiling students' learning. Adhering to the *Merdeka Curriculum* and international psychometric standards, SENOSTIK is a feasible technology-enhanced tool to assist teachers in generating adaptive instructions based on student characteristics. SENOSTIK's potential as a strategic tool to facilitate evidence-based assessments in Indonesian secondary schools has been proven. However, more research is required to expand its generalisation and explore the possibility of integrating it with intelligent technologies for a more in-depth and effective assessment analysis.

## ACKNOWLEDGEMENTS

The authors are especially thankful to SMA 13 Banjarmasin for their willingness to participate in the pilot test, and to educational departments and educators across Indonesia for their support. A special appreciation also goes to ILEIID and APB-LEAPS, organised by Akademi Pengajian Bahasa, Universiti Teknologi MARA, Malaysia, for providing SENOSTIK a platform for wider and international recognition.

## REFERENCES

1. Bautista-Díaz, M. L., Franco-Paredes, K., & Hickman-Rodríguez, H. (2022). Objetividad, validez y confiabilidad: atributos científicos de los instrumentos de medición. Educación y Salud Boletín Científico Instituto de Ciencias de La Salud Universidad Autónoma Del Estado de Hidalgo, 11(21), 66–71. https://doi.org/10.29057/icsa.v11i21.10048
2. Berrío García, N., & Zedán-Salinas, K. P. (2023). Calidad psicométrica de un instrumento: breve revisión. Poiésis, 45, 95–110. https://doi.org/10.21501/16920945.4522
3. Budiarti, R. P. N., & Kamila, F. I. R. (2024). Transformative Implementation of Android-Based Point of Sale System at Shafira's Buffet Stall. E3S Web of Conferences, 482, 02002. https://doi.org/10.1051/e3sconf/202448202002
4. Creswell, J. W., & Miller, D. L. (2000). Determining Validity in Qualitative Inquiry. Theory Into Practice, 39(3), 124–130. https://doi.org/10.1207/s15430421tip3903_2
5. Creswell, J. W., & Poth, C. N. (2016). Qualitative Inquiry Research Design: Choosing Among Five Approaches (H. Salmon, J. Scappini, C. Pearson, L. Larson, & M. Markanich, Eds.; Fourth Edi). Sage Publication.
6. Crossley, J. G. M. (2017). Assessing the Non-Cognitive Domains (pp. 348–372). https://doi.org/10.4018/978-1-5225-0531-0.ch018
7. Crump, T. G. (2024). Noncognitive Impact: Is There a Rationale for Assessing Postsecondary Students beyond Their Cognitive Skill Attainment? Creative Education, 15(02), 325–345. https://doi.org/10.4236/ce.2024.152020
8. Derilo, R. C. (2024). Synergistic Effect of Learning Environments, and Familial Factors on Generation Z Learners' Academic Achievement in Science. International Journal of Instruction, 17(4), 503–520. https://doi.org/10.29333/iji.2024.17428a
9. Dwi Saputra, R., Dinda Marsyanadya Putri, & Jadiaman Parhusip. (2024). Proporsi Individu yang Memiliki Telepon Genggam 2021–2023 dalam Perspektif Teknologi Informasi dan Komunikasi (TIK). Informatech : Jurnal Ilmiah Informatika Dan Komputer, 1(2), 183–188. https://doi.org/10.69533/2407sv12

10. González-Quezada, E., Soltero-Sánchez, A.L.A., Huerta-Chavez, I.A., & Figueroa-Ochoa, E. (2024). Diagnostic assessment of foundational knowledge and its relationship with the teaching and learning process in a university environment. Revista Gestión Universitaria.
11. Hadiastriani, Y., Djarot, P., & Chemistry, V. I. (2024). BIBLIOMETRIC INSIGHTS INTO ASSESSMENT PRACTICES IN THE MERDEKA CURRICULUM: IDENTIFYING OPPORTUNITIES FOR NEW ASSESSMENT DEVELOPMENT. JOURNAL OF SCIENCE EDUCATION AND PRACTICE, 8(1), 14–38. https://doi.org/10.33751/jsep.v8i1.9652
12. Huang, F., Mislevy, J., Wang, S., Wei, X., & Zhang, X. (2022). Editorial: Rigorous and high-quality efficacy studies of educational technology interventions. Frontiers in Education, 7. https://doi.org/10.3389/feduc.2022.968708
13. Isnawati, I. (2023). EFL TEACHERS' ASSESSMENT LITERACY. LLT Journal: A Journal on Language and Language Teaching, 26(2), 760–769. https://doi.org/10.24071/llt.v26i2.3654
14. Jannah, M., Maryani, I., & Santosa, A. B. (2024). Kesiapan Guru Sekolah Dasar dalam Implementasi Asesmen Diagnostik untuk Mendukung Pembelajaran Berdiferensiasi. Ideguru: Jurnal Karya Ilmiah Guru, 10(1), 451–459. https://doi.org/10.51169/ideguru.v10i1.1467
15. Junaidah, Ambiyar, & Fahmi Rizal. (2024). Mobile-Based Digital Assessment Transforming the Learning Support Efficiency. International Journal of Interactive Mobile Technologies (IJIM), 18(23), 44–56. https://doi.org/10.3991/ijim.v18i23.51339
16. Kusmawan, U. (2024). Transforming digital learning and assessment strategies in higher education. Multidisciplinary Reviews, 8(1), 2025016. https://doi.org/10.31893/multirev.2025016
17. Lisna, L., Pristiwaluyo, T., & Suhardi, I. (2024). Pengembangan Instrumen Asesmen Diagnostik Non Kognitif untuk Siswa SMP. Jurnal Basicedu, 8(5), 3529–3537. https://doi.org/10.31004/basicedu.v8i5.8183
18. Maufiroh, U. (2025). Sistematic Literature Review: Penerapan Asesmen Diagnostik Dalam Kurikulum Merdeka. COMSERVA : Jurnal Penelitian Dan Pengabdian Masyarakat, 4(9), 3010–3013. https://doi.org/10.59141/comserva.v4i9.2776
19. Molnár, G., & Kocsis, Á. (2024). Cognitive and non-cognitive predictors of academic success in higher education: a large-scale longitudinal study. Studies in Higher Education, 49(9), 1610–1624. https://doi.org/10.1080/03075079.2023.2271513
20. Musah, M. B. (2024). Assessing the Unassessable (pp. 248–266). https://doi.org/10.4018/979-83693-0880-6.ch017
21. O. Rulida, D., C. Cano, J., & R. Andrin, G. (2024). Non-Cognitive Skills as Correlates to Academic Performance among Senior High School Students. International Journal of Science and Management Studies (IJSMS), 222–249. https://doi.org/10.51386/25815946/ijsms-v7i1p131
22. Roncero, C. (2015). La validación de instrumentos psicométricos: un asunto capital en la salud mental. Salud Mental, 38(4), 235–236. https://doi.org/10.17711/SM.0185-3325.2015.032
23. Rosmiati, M., Maulid, H., Brotoharsono, T., Fathan Akbar, M., Lucky P, B., & Defari, W. (2023). APLIKASI CHECKBUN SEBAGAI MONITORING KESEHATAN IBU HAMIL BERBASIS ANDROID. Prosiding Konferensi Nasional Pengabdian Kepada Masyarakat Dan Corporate Social Responsibility (PKM-CSR), 6, 1–12. https://doi.org/10.37695/pkmcsr.v6i0.2108
24. Tikno, Dharmawan, Y. S., & Ngatini. (2024). Investigating Consumer Acceptance of Mobile Payment Services in Indonesia. Procedia Computer Science, 234, 1095–1102. https://doi.org/10.1016/j.procs.2024.03.104
25. Vitoratou, S., Uglik-Marucha, N., Hayes, C., & Pickles, A. (2023). An introductory comprehensive guide for assessing measurement tool quality: the Contemporary Psychometrics (ConPsy) Checklist. https://doi.org/10.31234/osf.io/t2pbj
26. Wu, J. R., Iwanaga, K., Qisti, D. A., Bezyak, J., Chan, F., & Tansey, T. N. (2025). Psychometric validation of the Motivation To Use Digital Tools At Work Scale in a sample of state vocational rehabilitation counselors. Journal of Vocational Rehabilitation, 62(1), 88–96. https://doi.org/10.1177/10522263241302222
27. Yansa, H., & Retnawati, H. (2021). Identifikasi Praktik dan Hambatan Guru dalam Asesmen Kognitif Matematika di Masa Pandemi COVID-19. Jurnal Elemen, 7(1), 84–97. https://doi.org/10.29408/jel.v7i1.2585