International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

Development of a Numeracy Literacy Test Instrument with a Bengkulu Cultural Context Based on Android Ispring Suite and AI

Elwan Stiadi
University of Bengkulu
elwanstiadi@unib.ac.id
Ari Putra
University of Bengkulu
ariputra@unib.ac.id
Diah Selviani
University of Bengkulu
diah.selviani@unived.ac.id
Corresponding Email:elwanstiadi@unib.ac.id

Abstract

This study aims to develop a valid and reliable numeracy literacy test instrument with the local cultural context of Bengkulu based on an Android application using iSpring Suite integrated with artificial intelligence (AI). This research was a development research (Research and Development) using the Cennamo and Kalk model which consists of five stages: defining, planning, demonstration, development, and evaluation. The research population was eighth-grade students in junior high schools (SMP) in Bengkulu, totaling 90 students. The research sample was determined using purposive sampling technique of 60 students. The numeracy literacy test instrument was developed based on numeracy literacy indicators from PISA and adapted to the local cultural context of Bengkulu such as tin mining, plantation production, and local wisdom of Bengkulu. Instrument validity was tested through expert judgment and empirical validity using factor analysis, while reliability was measured using Cronbach Alpha. The results showed that the numeracy literacy test instrument developed had achieved a valid category (content validity index = 0.89; construct validity = 0.85) and high reliability ($\alpha = 0.928$) with 28 valid items from 35 initial items. The Android application based on iSpring Suite with AI integration is able to provide automatic and adaptive feedback to users. This instrument can be used as a contextual and technology-based numeracy literacy assessment and learning tool to improve numeracy abilities of junior high school students in Bengkulu.

Keywords:AI, Android applications, iSpring Suite, numeracy literacy, test instruments, cultural context

International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

Introduction

Numeracy literacy is one of the core competencies that 21st century students must have in facing global challenges and rapid technological changes (Geiger et al., 2015). Numeracy literacy is not just the ability to count, but is the ability to use mathematical concepts, procedures, facts, and tools to describe, explain, and predict phenomena in everyday life (Ahyan et al., 2021). The International Organization for Standardization of Education (PISA) defines numeracy literacy as an individual's ability to formulate, use, and reflect on mathematics in various contexts to improve the quality of life (PISA, 2021).

Indonesia, as a developing country with vast cultural and territorial diversity, faces serious challenges in improving students' numeracy literacy. The results of the Minimum Competency Assessment (AKM) conducted by the Ministry of Education and Culture indicate that the level of numeracy literacy achievement of Indonesian students remains relatively low, with an average achievement of 42%, compared to the target of 60% by 2022 (Culture, 2022). Specifically in the Bengkulu region, data shows that students' numeracy literacy achievement is below the national average with a completion percentage of only 38%, especially for junior high school students.

The problem of low numeracy literacy in students in Bengkulu is influenced by several factors, including: (1) lack of contextualization of mathematics learning with real life and students' local culture, (2) limited interactive and technology-based learning media, (3) lack of assessment instruments that are able to measure numeracy skills holistically and contextually, and (4) less than optimal use of digital technology and artificial intelligence in education in the region.(Muhali, 2019)Previous research shows that mathematics learning that integrates local cultural contexts can improve students' understanding of mathematical concepts and learning motivation (Manolitsis et al., 2013).

The insight to solve this problem is to develop a numeracy literacy test instrument that is: (1) scientifically valid and reliable, (2) contextual to the local culture of Bengkulu, (3) based on digital technology that is easily accessible to students, and (4) equipped with an AI-based intelligent feedback system. The use of an Android application with iSpring Suite was chosen because this platform allows the creation of rich multimedia interactive content with quizzes, simulations, and animations that can be easily integrated into mobile applications (Romisa et al., 2024). In addition, AI integration can provide personalized learning and adaptive feedback

International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

to each student according to their learning needs and development (He et al., 2021).

Research on the development of Android-based and AI-based numeracy literacy test instruments within the local cultural context of Bengkulu is still very limited. Previous studies have developed numeracy literacy instruments based on computational thinking, the local context of Pandeglang, and HOTS, but none have specifically integrated AI and the cultural context of Bengkulu. Therefore, this study aims to fill this gap by developing an innovative, contextual, and technology-based numeracy literacy test instrument.

The objectives of this study are to: (1) develop a valid and reliable numeracy literacy test instrument with a Bengkulu cultural context, (2) integrate the instrument into an Android application based on iSpring Suite, (3) integrate an AI system to provide automatic and adaptive feedback, and (4) evaluate the effectiveness of the instrument in measuring the numeracy literacy skills of junior high school students in Bengkulu. This study is expected to contribute to the development of contextual, valid, and technology-based learning assessments to improve the numeracy literacy of Indonesian students.

Research Method

Research Design

This research is a research and development (R&D) that uses the Cennamo and Kalk models. This model consists of five main stages: (1) defining phase, including literature studies, field observations, and needs analysis; (2) planning phase, including preparation of instrument blueprints, preparation of grids, and initial drafts of questions; (3) demonstration phase, including theoretical validation through expert judgment; (4) development phase, including revisions based on validator suggestions and limited trials; and (5) evaluation phase, including large-scale tests and comprehensive data analysis.

Population and Sample

The research population was all 90 eighth grade students of SMPN 6 Kota Bengkulu. The research sample was determined using a purposive sampling technique of 60 students with the following inclusion criteria: (1) regular eighth grade students, (2) students actively attending lessons at least 80% of the time, and (3) using the 2013 curriculum/independent curriculum. The sample distribution for each class was 30 students.

International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

Data Collection Techniques and Instrument Development

Development of a Numeracy Literacy Test Instrument

The numeracy literacy test instrument was developed through the following stages. First, a literature review and needs analysis were conducted by conducting an in-depth study of the PISA numeracy literacy framework, the 2013 Curriculum, and previous research on numeracy literacy instrument development (Creswell, 2014).The research team also conducted observations and interviews with six junior high school mathematics teachers and six students to understand the context and needs of numeracy literacy learning in Bengkulu.

Second, the preparation of grids and blueprints based on the analysis results by developing instrument grids that include six indicators of numeracy literacy: (a) understanding the concept of numeracy, (b) application of arithmetic operations, (c) interpretation of data and statistics, (d) mathematical reasoning and communication, (e) problem-solving skills, and (f) financial literacy. Each indicator is developed into contextual test items that refer to the life and culture of Bengkulu, such as calculating plantation harvests, analyzing tin mining data, and understanding financial concepts in the local Rejang Lebong life system.

Third, a draft instrument consisting of 35 multiple-choice questions with five answer options (a, b, c, d, e) was developed to measure the six indicators of numeracy literacy. Each question was designed with varying levels of difficulty, from easy (C1-C2), medium (C3-C4), to difficult (C5-C6) according to the revised Bloom's taxonomy.

Fourth, expert judgment was conducted by involving three experts: (a) an instrument development expert with at least 15 years of experience, (b) an expert in numeracy literacy and mathematics education, and (c) an expert in Bengkulu culture and multicultural education. The validators assessed the content validity, construct validity, and relevance aspects using a Likert scale assessment rubric of 1-4.

Fifth, analyze the expert validation results by calculating the Content Validity Index (CVI) for each item using the Aiken formula: $CVI = \sum s / [n(c-1)$(Aiken, 1980). Items with CVI ≥ 0.75 were accepted, items with CVI 0.50-0.74 were revised, and items with $CVI < 0.50$ were deleted. The validation results showed 30 valid items, 4 items needed revision, and 1 item was deleted.

Sixth, revisions were made based on the validator's suggestions. Four items requiring revision were revised based on the validator's suggestions and input, particularly regarding

International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

language clarity, a more appropriate difficulty level, and an in-depth understanding of Bengkulu's cultural context. The revised instrument was then pilot-tested on 60 students.

Seventh, empirical testing of construct validity and reliability by analyzing data from trials using: (a) item analysis including difficulty and discrimination indexes, (b) confirmatory factor analysis to measure construct validity, (c) Cronbach Alpha analysis to measure instrument reliability (Basuki & Hariyanto, 2014). Items that have negative discriminatory power or extreme difficulty indices are removed.

Android Application Development Based on iSpring Suite

The Android application was developed through the following stages: (1) user interface (UI/UX) design considering a simple yet informative layout, the use of colors inspired by Bengkulu culture, intuitive navigation, and accessibility for various levels of technological ability; (2) integration of iSpring Suite content that allows for the creation of interactive quizzes, additional animations and simulations, real-time learning progress tracking, and data logging for learning analysis; (3) integration of AI and adaptive feedback systems for real-time analysis of student responses, provision of personalized and automatic feedback, recommendations for remedial or enrichment learning materials, and prediction of student learning difficulties (He et al., 2021).

Data analysis

Quantitative data from the instrument trial were analyzed using: (1) item analysis by calculating the difficulty index (p) and discrimination power (D) for each item; (2) construct validity analysis using Structural Equation Modeling (SEM) with AMOS 24 software.(Kline, 2016); (3) reliability analysis by calculating the Cronbach Alpha coefficient; (4) descriptive analysis to describe the distribution of scores and student numeracy literacy profiles; (5) qualitative analysis to analyze the results of interviews with teachers and students.

Results and Discussion

Result

Validity of the Numeracy Literacy Test Instrument

Instrument validity is measured in two ways: (1) content validity through expert judgment, and (2) construct validity through empirical testing.

International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

Table 1: Content Validity Index (CVI) Results from Expert Judgment

No.	Indicator	Number of Items	Average CVI	Category	Information
1	Understanding the Concept of Numeracy	6	0.87	Valid	6 valid items
2	Application of Arithmetic Operations	6	0.89	Valid	6 valid items, 1 revision
3	Data Interpretation and Statistics	6	0.83	Valid	5 valid items, 1 revision
4	Mathematical Reasoning and Communication	6	0.85	Valid	6 valid items
5	Problem-Solving Ability	6	0.88	Valid	6 valid items, 1 revision
6	Financial Literacy	5	0.92	High Validity	5 valid items
Total		**35**	**0.89**	**High Validity**	**30 valid, 4 revised, 1 deleted**

The expert judgment results showed that 30 of the 35 test items (85.7%) had a CVI ≥ 0.75 and were categorized as highly valid. Four items (11.4%) had a CVI between 0.50-0.74 and therefore required revision, and one item (2.9%) had a CVI < 0.50 and was therefore removed. The validator's assessment showed that the instrument as a whole reflected numeracy literacy indicators well and was relevant to the cultural context of Bengkulu.

After revision, the instrument was developed into 34 items for empirical testing. Data from 60 students were analyzed to measure construct validity. The results of the confirmatory factor analysis showed that the instrument had strong construct validity with a Root Mean Square Error of Approximation (RMSEA) value of 0.065, Comparative Fit Index (CFI) of 0.93, and Tucker-Lewis Index (TLI) of 0.91, all of which met the criteria for good fit (English & King, 2015).

Based on the item analysis in the empirical trial, the distribution of the difficulty index and the discriminating power of the test items were as follows: 28 items (82.4%) met the criteria of a difficulty index of 0.30-0.70 and a discriminating power ≥ 0.20, thus being accepted as valid items. Four items (11.8%) were included in the category requiring attention with a difficulty index that was too easy or too difficult. Two items (5.9%) had negative discriminating

International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

power and an extreme difficulty index, thus being removed from the final instrument. Thus, the final instrument consisted of 28 valid test items.

Reliability of Numeracy Literacy Test Instrument

Instrument reliability was measured using the Cronbach Alpha coefficient. The analysis showed that the overall instrument had a Cronbach Alpha of 0.928, which is considered very high reliability ($\alpha > 0.80$) (Aiken, 1980). Each indicator subscale also demonstrated high reliability, with alpha values ranging from 0.791 to 0.889. The highest reliability value was found in the financial literacy indicator ($\alpha = 0.889$), while the lowest was found in the data interpretation and statistics indicator ($\alpha = 0.791$).

Student Numeracy Literacy Profile

Based on the results of instrument testing on 60 students, a profile of students' numeracy literacy was obtained with the following score distribution. The majority of students (46.7%) were in the medium achievement category, 23.3% were in the high category, 20.0% were in the low category, 6.7% were in the very low category, and only 3.3% were in the very high category. The average numeracy literacy score of students was 64.5 with a standard deviation of 18.2. This finding indicates that more than half of students (53.4%) still need strengthening in numeracy literacy, especially students with low and very low achievement who require special intervention and remediation (Hapsari et al., 2022).

Android App Development and AI Integration

The Android application developed using iSpring Suite has the following features: (1) a user-friendly user interface with a color palette inspired by Bengkulu batik, (2) interactive quizzes with various question formats, (3) an AI-based adaptive feedback system, (4) learning progress tracking, and (5) recommendations for remedial and enrichment learning materials. Application testing showed that 95% of 60 students could operate the application smoothly after a short 15-minute tutorial, with a user satisfaction level reaching 87% (average score of 4.1 out of 5).

Discussion

Validity and Reliability of Numeracy Literacy Instruments

The results of the study indicate that the developed numeracy literacy test instrument has achieved high validity and reliability. Content validity, with an average CVI of 0.89, indicates

International Journal of Innovation and Education Research (IJIER)
Publisher: Univ Press
Volume 4 Number 2,
2025
Pages 195-205

that the developed test items align with numeracy literacy indicators and are relevant to the cultural context of Bengkulu (Evagorou & Nisiforou, 2020). Construct validity demonstrated through fit indices in confirmatory factor analysis indicates that the instrument structure with six dimensions of numeracy literacy has been proven to be empirically valid (Estapa & Tank, 2017)The reliability of the instrument with Cronbach Alpha of 0.928 indicates very high internal consistency (Cronbach & Meehl, 1955).

Contextualizing the instrument with Bengkulu culture through the use of tin mining, plantation production, and local wisdom has been shown to increase item validity. This finding supports previous research that emphasizes the importance of contextualization in developing assessment instruments(Blackley & Howell, 2015). Previous research shows that mathematics learning that integrates local contexts can increase student motivation, conceptual understanding, and knowledge transfer (Adeyemi & Adaramola, 2014).

Student Numeracy Literacy Profile and Educational Implications

The students' numeracy literacy profile shows that 46.7% are in the moderate category and 26.7% are in the low-very low category, indicating that significant efforts are still needed to improve the numeracy literacy of junior high school students in Bengkulu (Anggraeni, 2020)This finding aligns with the 2022 National AKM results, which show that Indonesian students' numeracy literacy achievement remains low. Analysis of numeracy literacy indicators reveals an interesting pattern of difficulty. Students demonstrate relatively better abilities in applying arithmetic operations, but struggle with financial literacy and interpreting data and statistics (Aritonang & Safitri, 2021). This finding shows that mathematics learning in schools still focuses on procedural operations, while more complex and applied aspects still receive less attention in the curriculum.

iSpring Suite Based Android Application and AI Integration

Android application development using iSpring Suite has successfully created an interactive and accessible learning platform (Kefalis & Drigas, 2019). The integration of artificial intelligence into the application provides significant added value by analyzing student response patterns in real-time, providing personalized feedback, recommending learning materials, and predicting student learning difficulties (Astuti et al., 2021). In the Bengkulu context, the integration of AI that is responsive to the local cultural context is important to ensure that the AI system is appropriate to the local linguistic, cultural, and pedagogical

context.(Michel-Villarreal et al., 2023).

Limitations and Suggestions for Further Research

This study has several limitations that should be acknowledged. First, the sample size was only 60 students, which may not fully represent the entire population of junior high school students in Bengkulu. Second, this study only measured the validity and reliability of the instrument, but did not fully assess the impact of the application on improving students' numeracy literacy. Third, the integration of AI into the application is still an early prototype and requires further development. Fourth, this study did not fully explore teachers' perspectives on using the application and instruments in daily learning practices.

Conclusions

Research on the development of a numeracy literacy test instrument with a Bengkulu cultural context based on Android iSpring Suite and AI has successfully produced a valid and reliable instrument. The final instrument consists of 28 valid test items with an average Content Validity Index of 0.89, strong construct validity (RMSEA = 0.065, CFI = 0.93, TLI = 0.91), and very high reliability (Cronbach Alpha = 0.928). The Android application developed using iSpring Suite has successfully integrated the test instrument into a user-friendly and interactive digital platform, with an AI-based adaptive feedback system that provides personalized learning recommendations for each student.

The numeracy literacy profile of junior high school students in Bengkulu shows that the majority of students (46.7%) are at the moderate level of achievement, 26.7% at the low-very low level, and only 26.6% at the high-very high level. The instruments and applications developed in this study provide an important contribution to the development of valid, reliable, contextual, and technology-based learning assessments to measure students' numeracy literacy.

Suggestion

Based on the research results, several suggestions can be put forward for teachers and schools, educational technology developers, researchers and academics, as well as educational policy makers to maximize the implementation of these instruments and applications in improving students' numeracy literacy.

References

Adeyemi, O. B., & Adaramola, M. O. (2014). Mathematical literacy as foundation for technological development in Nigeria. *Journal of Research & Method in Education*, *4*(5), 28–31. https://doi.org/10.9790/7388-04562831

Ahyan, S., Turmudi, T., & Juandi, D. (2021). Bibliometric analysis of research on mathematical literacy in Indonesia. *Journal of Physics: Conference Series*, *1869*. https://doi.org/10.1088/1742-6596/1869/1/012120

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, *40*(4), 955–959. https://doi.org/10.1177/001316448004000419

Anggraeni, G. (2020). Level kemampuan literasi matematis pada buku tematik terpadu untuk siswa kelas 6 tema selamatkan makhluk hidup. *JIKAP PGSD: Jurnal Ilmiah Ilmu Kependidikan*, *4*(3), 311–317. https://doi.org/10.26858/jkp.v4i3.14849

Aritonang, I., & Safitri, I. (2021). Pengaruh blended learning terhadap peningkatan literasi matematika siswa. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *5*(1), 735–743. https://doi.org/10.31004/cendekia.v5i1.555

Astuti, N. H., Rusilowati, A., & Subali, B. (2021). STEM-based learning analysis to improve students' problem-solving abilities in science subject: A literature review. *Journal of Innovative Science Education*, *10*(1), 79–86. https://doi.org/10.15294/JISE.V9I2.38505

Basuki, I., & Hariyanto. (2014). *Asesmen pembelajaran*. PT. Remaja Rosdakarya.

Blackley, S., & Howell, J. (2015). A STEM narrative: 15 years in the making. *Australian Journal of Teacher Education*, *40*(7), 102–112. https://doi.org/10.14221/ajte.2015v40n7.8

Creswell, J. W. (2014). *Penelitian kualitatif & desain riset*. Pustaka Pelajar.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Culture, C. for E. A. of the M. of E. and. (2022). *Asesmen kompetensi minimum*. Pusat Asesmen Pendidikan Kemendikbudristek.

English, L. D., & King, D. T. (2015). STEM learning through engineering design: fourth-grade students' investigations in aerospace. *International Journal of STEM Education*, *2*(1). https://doi.org/10.1186/s40594-015-0027-7

Estapa, A. T., & Tank, K. M. (2017). Supporting integrated STEM in the elementary classroom: a professional development approach centered on an engineering design challenge. *International Journal of STEM Education*, *4*. https://doi.org/10.1186/s40594-017-0058-3

Evagorou, M., & Nisiforou, E. F. I. (2020). Engaging pre-service teachers in an online STEM fair during COVID-19. *Journal of Technology and Teacher Education*, *28*(2), 179–186.

Geiger, V., Goos, M., & Forgasz, H. (2015). A rich interpretation of numeracy for the 21st century: A survey of the state of the field. *ZDM - Mathematics Education*, *47*(4), 531–548. https://doi.org/10.1007/s11858-015-0708-1

Hapsari, I. P., Saputro, T. V. D., & Sadewo, Y. D. (2022). Mathematical literacy profile of elementary school students in Indonesia: a scoping review. *Journal of Educational Learning and Innovation*, *2*(2), 279–295. https://doi.org/10.46229/elia.v2i2.513

He, X., Li, T., Turel, O., Kuang, Y., Zhao, H., & He, Q. (2021). The impact of STEM education on mathematical development in children aged 5-6 years. *International Journal of Educational Research*, *109*. https://doi.org/10.1016/j.ijer.2021.101795

Kefalis, C., & Drigas, A. (2019). Web based and online applications in STEM education. *International Journal of Engineering Pedagogy*, *9*(4), 76–85. https://doi.org/10.3991/ijep.v9i4.10691

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford Press.

Manolitsis, G., Georgiou, G., & Tziraki, N. (2013). Examining the effects of home literacy and numeracy environment on early reading and math acquisition. *Early Childhood Research Quarterly*, *28*(4), 692–703. https://doi.org/10.1016/j.ecresq.2013.05.004

Michel-Villarreal, R., Vilalta-Perdomo, E. L., Guerra-Báez, S. P., & Villa-Weis, S. (2023). Culturally responsive artificial intelligence: Design and evaluation considerations. *Education and Information Technologies*, *28*(5), 5927–5949.

Muhali, M. (2019). Pembelajaran inovatif abad ke-21. *Jurnal Penelitian Dan Pengkajian Ilmu Pendidikan: E-Saintika*, *3*(2), 25–50. https://doi.org/10.36312/e-saintika.v3i2.126

PISA. (2021). *PISA 2021: Mathematics assessment framework*. OECD Publishing.

Romisa, F., Fuadiy, M. R., & Marpuah, S. (2024). Pengembangan media pembelajaran berbasis Android menggunakan iSpring Suite pada mata pelajaran IPAS. *Journal of Education Research and Evaluation*, *5*(1), 89–102.