Journal of Applied Science, Engineering, Technology, and Education *ISSN: 2685-0591 (online)*
Vol. 7 No. 1 (2025)
https://doi.org/10.35877/454RI.asci3785

Paper vs. Digital Assessments: Evaluating Critical Thinking on Ecological Issues in Indonesian Madrasahs

Zainal Berlian[a,*], Muhammad Isnaini[a], Ratna Farwati[b], Kartika Metafisika[c], Yessy Velina[d], & Viona Fitriyanni[b]

[a]*Islamic Education Management Postgraduate Study Program, Universitas Islam Negeri Raden Fatah Palembang, Indonesia*
[b]*Chemistry Education Undergradute Study Program, Universitas Islam Negeri Raden Fatah Palembang, Indonesia*
[c]*Madrasah Ibtidaiyah Teacher Education Undergradute Study Program, Sekolah Tinggi Agama Islam Taruna Surabaya, Indonesia*
[d]*Earth, Life, and Environment Postgraduate Study Program, Kyoto University, Japan*

Abstract

This research aims to determine the differences in the results of assessments of critical thinking skills of State Madrasah Ibtidaiyah (MIN) students or State Islamic schools in Indonesia using paper-based and digital-based tests. Data was collected from 314 participants representing Madrasas on four large islands in Indonesia, namely MIN Bogor Regency, West Java, MIN Sinjai Regency, South Sulawesi, MIN Palembang City, South Sumatra, and MIN Balikpapan City, East Kalimantan. The research method used a quantitative survey. A total of 24 valid Critical Thinking Skills Items with the Facione Critical Thinking Skills indicator as an instrument for measuring critical thinking skills. Factors such as motivation, technological accessibility, and effectiveness of test formats contribute to differences in results. This study provides new insights into the effectiveness of technology-based assessment in Madrasa education, especially in evaluating critical thinking skills. These findings lead to policy recommendations for improving the Madrasa assessment system in the digital era.

Keywords: assessment; critical thinking skills; elementary education; madrasah ibtidaiyah.

Received: 19 March 2025 Revised: 27 April 2025 Accepted: 29 April 2025

1. Introduction

Today and in the future, knowledge acquisition will no longer be limited by accessibility but by selection, attention, and ability to manage information (Christensen & Bailey, 1997; Kisilowska-Szurmińska, 2023). In the current era of social media, anyone can access much information, but it is not necessarily useful in their lives and even leads to misinformation (Bainiashova, 2023; Kilickaya & Ersoy, 2016). It requires a person's ability to think slowly. According to Kahneman, regarding the dual system thinking theory, a person needs to have the ability to think slowly or critically to hone their thinking skills and reflect on their knowledge for their life (Halpern, 2013; Kannengiesser & Gero, 2019). Critical thinking skills equip students to face advances in information technology and support education that leads to Sustainable Development Goals (Wani & Hussian, 2024). These critical thinking skills must be taught and assessed since elementary school to get maximum benefits (Ennis, 1989; Kennedy et al., 1991).

Even though it is important, Indonesian students' critical thinking skills are still relatively low. It aligns with findings that students struggle with basic classification, assessing information, drawing conclusions, and applying problem-solving models (Ferrary et al., 2023). A study of the critical thinking skills of fourth-grade elementary school students in Sidoarjo Regency in learning about Energy Resources shows that students still have difficulty concluding and providing strong reasons to overcome their answers (Fitanti et al., 2024). Based on the results of identifying factors that cause the low critical thinking skills of elementary school students, from the student side, it was found that students often give unstructured answers, misunderstand questions, have misunderstandings about the subject matter,

* Corresponding author.

Email address: zainalberlian68@radenfatah.ac.id

and tend to rely on memorization rather than deep understanding. These factors show that students struggle to apply critical thinking skills to solve problems (Sarwanto et al., 2021). This difficulty can be influenced by a lack of learning practices in the classroom to develop critical thinking skills (Lin et al., 2017). Apart from that, it is known that male students' socio-economic conditions and critical thinking skills tend to be higher than those of female students (Gelerstein et al., 2016; Kawuryan et al., 2022). Research on improving critical thinking skills in elementary schools in Indonesia has been widely conducted (Anggraeni et al., 2023). However, no studies have specifically examined the factors influencing the results of critical thinking skills assessments.

In today's digital era, assessments in measuring critical thinking skills using digital media provide extraordinary benefits compared to paper-based ones because they can provide feedback directly to students, can be adapted to students' abilities, and provide attraction and involvement (Hussein et al., 2019; Shute & Rahimi, 2017; Utaminingsih et al., 2024). Various assessment platforms such as Kahoot and Quiziss that are ready to be used by teachers function as evaluation tools and can influence students' motivation, engagement and self-efficacy (Maraza-Quispe, 2024; Shyr et al., 2021). Students can also easily adapt to Android applications for learning and assessment. It is known that when using Android applications at school, students are only briefly introduced to the application. They use the application independently, and from the assessment results, it is known that it can improve students' critical thinking abilities (Isrokatun et al., 2022).

The Ministry of Religious Affairs (MORA) oversees Madrasah Ibtidaiyah, an Islamic-based school in Indonesia. It seeks to improve thinking skills by increasing student literacy and numeracy through the Indonesian Madrasah Competency Assessment (AKMI) program. AKMI program is a collaboration between MORA and the World Bank in the Realizing Education's Promise - Madrasah Education program Quality Reform (REP-MEQR) 2022-2024. One of the program's objectives is to develop and standardize an assessment system for primary school students (WorldBank, 2024). The AKMI program also involves madrasa teachers throughout Indonesia to develop instruments, develop training modules, and train teachers throughout Indonesia online and offline (Widhiarso & Ridho, 2022). The level of competency achievement from the AKMI assessment results is a stepping stone for Madrasas to improve the learning process and provide facilities and infrastructure according to learning needs. One of the advanced socialization materials from AKMI is an effort to improve learning by implementing various classroom learning strategies, such as problem-based learning, project learning, inquiry, and discovery learning. It is hoped to increase students' literacy and numeracy (Millah et al., 2022; Rosa, 2022; Widhiarso & Ridho, 2022). In assessment practice, AKMI is implemented on a computer basis, although in daily learning practice, summative and formative assessments are still paper-based assessments.

This difference can raise the question of whether the difference between paper and digital formats can affect the assessment results. If assessment is used in digital format, using a device to support learning is the easiest and cheapest. However, despite the extraordinary benefits of using digital assessment in the learning process, the obstacles to its use are the facilities and infrastructure and the teacher's attitude towards using devices in the classroom. The use of devices in the learning process in the classroom, especially in basic education, still raises debate. Even though it has a positive impact, namely facilitating feedback from formative assessments, facilitating independent and peer learning, and increasing student involvement and motivation, it also has a negative side because it can increase dependence on devices, increase distraction, and reduce social skills (Bacak et al., 2022; Ciampa & Gallagher, 2013).

To answer this question, researchers want to know the differences in the results of assessments of critical thinking skills on ecological issues for elementary school students between paper and digital formats. Critical thinking skills are the focus of measurement because researchers want to know the follow-up of AKMI results, namely scientific literacy skills. It is known that critical thinking skills in science learning are associated with a person's knowledge regarding understanding the nature of science (NOS), which has a positive relationship with scientific literacy (Almeida et al., 2023; Primasari et al., 2020; Ridzal & Haswan, 2023). Critical thinking skills are also part of a person's basic literacy (Vieira & Tenreiro-Vieira, 2016). Apart from that, the application of a constructive learning approach or model in AKMI follow-up training for teachers, namely Problem-Based Learning, Project-Based Learning, Inquiry, And Discovery Learning can improve students' critical thinking skills (Anggraeni et al., 2023; Manurung & Pappachan, 2025; Sumarni & Kadarwati, 2020; Sutiani et al., 2021; Ubaidillah et al., 2023). Ecological problems are raised in the assessment of critical thinking skills because they are issues of sustainable development goals which need to be introduced from a basic level to students. Ecology is also closely related to scientific literacy competencies so that students can reflect themselves as members of society by utilizing scientific knowledge and processes and is part of the science and social studies learning outcomes of the National Primary School Curriculum grade 3rd the 4thth (Jeong et al., 2021; Queiruga-Dios et al., 2020; Shin-Jia & Tien-Chi, 2024).

No one has studied the test format factor in terms of the results of measuring critical thinking skills. Previous research results Brüggemann et al. (2023) examined the impact of various test formats (pen and paper, computer-based, and computer-adaptive) on test anxiety and motivation among German fourth-grade students during a standardized reading comprehension test involving 387 participants aged 9 -10 years. The results showed no difference in anxiety between exam formats. However, the digital format provided higher motivation at the start than the paper format, although this motivation decreased during the exam. Apart from motivation and anxiety, it is also known that there is no significant difference in cognitive load between the two formats (Brüggemann et al., 2023a). Perry et al. (2022) conducted a literature review study of 20 articles regarding the use of computer-based assessments (CBA), which showed that students preferred CBA to paper-based; however, practice was needed to adapt to the digital format when implementing the exam. Thus, this research aims to discover whether the difference in test formats (pen- and paper-based and digital-based tests) influences the results of students' critical thinking skills assessments.

1.1. Madrasas and Their Challenges

Indonesia is the country with the largest Muslim population in the world, with 90% of its 260 million population embracing Islam amidst linguistic, ethnic, and religious diversity and democracy as a government system (Azra, 2021)**.** The large population of Indonesian Muslims and the interest of the Indonesian Muslim community in sending their children to Islamic educational institutions have made Madrasas and Islamic-based schools very popular among the public. Parents' desire to balance religious and non-religious education in their children's lives is a factor that influences middle-class Muslim communities to choose Islamic-based schools. Research shows that there is a total of 282 Non-Formal Islamic Schools (Madrasah Diniyah), 165 Elementary Schools (Madrasah Ibtidaiyah), 93 Middle Schools (Madrasah Tsanawiyah), and 47 High Schools (Madrasah Aliyah) in Indonesia. In total, there are 587 madrasas at various levels of education, most of which are concentrated on the island of Java, especially in West Java, and some in South Sumatra Province. Just like formal schools under the Ministry of Education, the education period at Madrasah Ibtidaiyah is 6 years, Madrasah Tsanawiyah and Madrasah Aaliyah is 3 years each. This number does not include private madrasas and Islamic boarding schools outside the Ministry of Religion, namely those under the Ministry of Education.

Some formal primary schools and secondary education institutions are under the auspices of the Ministry of Education, and some are under the Ministry of Religious Affairs. The speciality of the Madrasah is the Islamic religious curriculum which accompanies the general education curriculum, including lessons on Al-Qur'an Hadith, History of Islamic Culture, Aqidah Akhlak, Arabic and Fiqh and is thick with Islamic religion, local culture, and a family atmosphere. There are worship habituation programs such as congregational prayers, zakat, and fasting in Ramadan, which are part of the five pillars of Islam. Even madrasas also provide dormitory programs to build the character of Muslims who obey religious commands (Sofanudin et al., 2016). In contrast to Islamic-based schools under the Ministry of Education, Madrasahs under the Ministry of Religion have challenges in implementing their curriculum because they are tied to the Madrasah curriculum, which must be implemented strictly, especially in State Madrasahs. As mandated by the National Education System Law number 20 of 2003, the burden on Madrasas in implementing these two curricula is increasingly heavy because, since the Dutch colonial era, Madrasas were built specifically to study Islamic religious knowledge so that the religious curriculum cannot be easily reduced (Ihsan et al., 2024). Based on the Decree of the Ministry of Religion (KMA) number 450 of 2024, the division of study hours between general studies is 87% and religion is 13%. Teachers need strategies to design student-centred, holistic, and interdisciplinary learning and choose important material rather than just teaching book content and exercises in preparation for exams (Moslimany et al., 2024). Even though there are challenges in implementing these two curricula, Madrasah's obligation to provide general education can be a solution to prepare Madrasah students' skills in facing future employment choices. Apart from that, good character, which is instilled in habits such as discipline in worship, emphasis on morals as an important part of educational goals, respect, and tolerance for diversity, can strengthen the character of Madrasah graduates in facing the future.

Another challenge is that the existence of Madrasas under the Ministry of Religion creates funding disparities. The Ministry of Education's budget, which focuses on education, can provide diverse and sustainable programs and is decentralized. Meanwhile, the Ministry of Religion's budget focuses on education and is centralized. Hence, infrastructure development, professional development, and the procurement of learning tools and materials are not as well facilitated as institutions under the Ministry of Education (Kusnanto et al., 2023). It is known that State Madrasas are the choice of lower-middle-class people, while middle-class people choose to send their children to

Islamic-based private schools or full-day private Madrasas with various school programs (Mariana & Helmi, 2022; Martin, 2023; Park & Niyozov, 2008). With these limitations, various programs to increase teacher professionalism are implemented massively through Massive Open Online Courses (MOOC) on the website https://pintar.kemenag.go.id/, as well as teacher training programs through follow-up to AKMI results, which are carried out simultaneously throughout Indonesia (Millah et al., 2022; Zarkasy et al., 2023). With an allocation time percentage of 87% for general learning and teacher training through various sustainable platforms with various existing limitations, it is hoped that there will be an increase in student literacy every year.

1.2. Primary School Students Critical Thinking Skills.

Critical thinking skills are rooted in behavioural psychology, a higher-order thinking skill (Erwin, 2000). There is a difference of opinion regarding critical thinking between Ennis, who believes that critical thinking can be independent of the knowledge domain, and McPeck, who argues that critical thinking skills must be tied to a particular knowledge domain. Martin argues that critical thinking must be based on a moral compass and justice. Siegel emphasizes that the main thing in critical thinking is reasoning, regardless of whether it is tied to a particular knowledge domain (Mason, 2007). Reasoning needs to be understood and built with strategic and objective steps. Thus, reasoning is an important critical thinking skill (Mason, 2007). With critical thinking skills, students are expected to analyze, evaluate and synthesize information consciously and reflectively to make appropriate decisions and judgments amidst the rapid flow of information they can access easily (P. A. Facione & others, 2011).

Facione clearly distinguishes between critical thinking skills and dispositions (P. A. Facione & others, 2011). Critical thinking refers to an individual's inclination or propensity to engage in critical thinking, including open-mindedness, persistence, reflectivity, curiosity, and self-confidence. In contrast, critical thinking skills are the cognitive abilities used to analyze, evaluate, and synthesize information. Although dispositions influence the likelihood of using critical thinking skills, the two concepts are different; dispositions are about attitudes towards critical thinking, while skills related to the practical application of critical thinking processes that can be cultivated and accustomed are the ability to think critically through student-centred learning processes (P. Facione, 1990; Hitchcock, 2018; Xu et al., 2023).

According to Facione, experts formulate the cognitive skills that someone who thinks critically must have, namely:

1) Interpretation: The ability to understand and express the meaning or significance of various experiences, data, events, judgments, beliefs, etc., is a core cognitive skill in critical thinking.
2) Analysis: Critical thinking involves breaking down complex information into constituent parts, examining their relationships, and understanding underlying patterns and structures.
3) Evaluation: Critical thinking requires assessing the credibility, relevance, and quality of information, arguments, and evidence. It involves making judgments based on logical reasoning and evidence.
4) Inference: Critical thinking involves drawing logical conclusions and making reasonable predictions based on available information and evidence. It requires identifying implicit assumptions and implications.
5) Explanation: Critical thinking includes providing clear and coherent justifications, reasons, and explanations for one's beliefs, decisions, and actions. It involves communicating ideas effectively and logically.
6) Self-regulation: Critical thinking includes the ability to monitor and control one's thought processes, including being aware of biases, assumptions, and emotions that may influence decision-making. It involves a reflective and self-correcting attitude (P. A. Facione & others, 2011).

Critical thinking skills with Facione indicators are widely used in research to measure critical thinking skills in science learning (Dewi et al., 2022; Hasanah et al., 2021; Hsu et al., 2022; Mastuti et al., 2022; Ubaidillah et al., 2023). In elementary school, research on critical thinking is more in mathematics and science in fifth grade (Juliyantika & Batubara, 2022). Hsu et al. (2022) measured critical thinking skills in their research using the indicators used by Facione by issuing the Self-Regulation Indicator because it is synthetic and refers to the five previous indicators.

2. Method

This research is quantitative research with a survey method. The sampling technique was by selecting representatives of State Madrasah Ibtidaiyah (MIN) purposively on large islands spread throughout Indonesia, namely MIN in Bogor Regency, West Java Province, on Java Island; MIN in Sinjai Regency, South Sulawesi Province, Sulawesi Island; MIN in Balikpapan City, West Kalimantan Province, Kalimantan Island; and MIN in Palembang City, South Sumatra Province, Sumatra Island. MIN Palembang and MIN Balikpapan are in the provincial capital with easy accessibility,

while MIN Bogor and MIN Sinjai are in the district. MIN Sinjai is in a remote location, while MIN Bogor is not far from the capital, Jakarta, so access and infrastructure are still accessible. Internet access is very limited at MIN Sinjai, with frequent rolling blackouts, whereas at other Madrasas, there is smooth internet network access. Not all madrasahs allow students to bring devices to the madrasah. The purpose of selecting samples purposively is to take MIN representatives from each island and from each island there are representatives of cities and districts. However, this method has the potential for selection bias because it is not selected randomly or stratified. The reason for using this sampling method is because of the effectiveness of accessibility in data collection.

The research was conducted in September 2024. The survey targeted 314 students in grades 3rd, 4th, and 5th. The value corresponds to the thinking ability of students starting to think at a higher level and is the target of the AKMI program (Gendenjamts, 2023; Hollett, 2022; WorldBank, 2024). Table 1 shows the distribution of respondents in each region.

Table 1. Distribution of Respondents According to Grade and MIN.

State Ibtidaiyah Madrasah (MIN)	Grade	Number of Respondents
MIN Bogor Regency	3rd	33
	4th	52
MIN Palembang City	3rd	30
	5th	27
MIN Sinjai Regency	3rd	15
	4th	19
	5th	11
MIN Balikpapan City	3 rd	52
	4 th	62
	5 th	13

Source: Primary Data

Table 2. Homogeneity and Normality Test Results.

Category	Lavene test P-value	Saphiro Wilk test P-value
MIN Bogor Regency, West Java	0.1974 (homogeneous)	0.3145 (normal)
MIN Palembang City, South Sumatra		0.01528 (abnormal)
MIN Sinjai Regency, South Sulawesi		0.3091 (normal)
MIN Balikpapan City, East Kalimantan		0.1849 (norrmal)

Source: Statistical Analysis from Primary Data.

The instrument was developed as a 4-choice multiple choice (a, b, c, and d) to make it easy for students to answer because it corresponds to the exam question format they encounter in their daily learning. The number of words for each question is adjusted to students' readability to reduce the readability factor of the questions and uses illustrations to make it easier for students to understand the questions ((Ishartono et al., 2021; Melawati et al., 2023). Of the 41 questions developed, referring to indicators of Facione's critical thinking skills (Interpretation, Analysis, Inference, Evaluation and Explanation) related to ecological issues, 24 items were declared valid by experts and practitioners, and after going through the validity test of the items using the Rasch Model, all of them were declared valid with the reliability value for the items was 0.89, which indicated that the quality of the items was good. Data collection was carried out at the same time duration, namely 1 hour, to work on 24 items, namely using Google form for digital format with BYOD (Bring Your Device) at MIN Bogor Regency and MIN Balikpapan City and paper format at MIN Sinjai Regency and MIN Palembang City. Consider using Google Forms because there is no element of gamification. Thus, the only influencing factor is the difference in the format of the assessment tool.

The results of the normality and homogeneity test of the data in Table 2 indicate that the data is not normal and homogeneous. Therefore, data analysis was carried out using the Kruskal-Wallis test to measure differences in data

from more than two independent data sets, followed by the Dunn test to show the significance of mean differences between groups using the Rstudio application. Boxplots were used to compare median positions between groups.

3. Results and Discussion

3.1. Results

From the assessment results in Table 3, it is known that students' critical thinking abilities in grades 3rd, 4th, and 5th The values for the sample are 0 to 96, with a mean value of 58 and a standard deviation of 20.35. From these scores, five ranges of very low, low, medium, high and very high scores were created in Table 3. Seen from Figure 3. A total of 29.83% were in the low and very low grades, 64% of the students were in the medium position, and 8.83% were in high and very high positions. From these data, it can be said that students' critical thinking abilities vary greatly, and the medium category scores have a very wide range; this shows that it is necessary to carry out a more in-depth analysis based on subgroups, namely based on Value and MIN.

Table 3. Descriptive Statistics Calculations.

Value	N Statistic	Minimum	Maximum	Mean	Std.Deviation
Occurs N	305	0	96	58	20.35

Source: Statistical calculations from primary data.

Table 4. Categorization of Students' Critical Thinking Abilities.

Score Range	**Category**
X ≤ 26.4	Very Low
26.40 < X ≤ 46.75	Low
46.75 < X ≤ 83.66	Everage
83.66 < X ≤ 89.5	High
89.50 > X	Very high

Source: Statistical calculations from primary data.

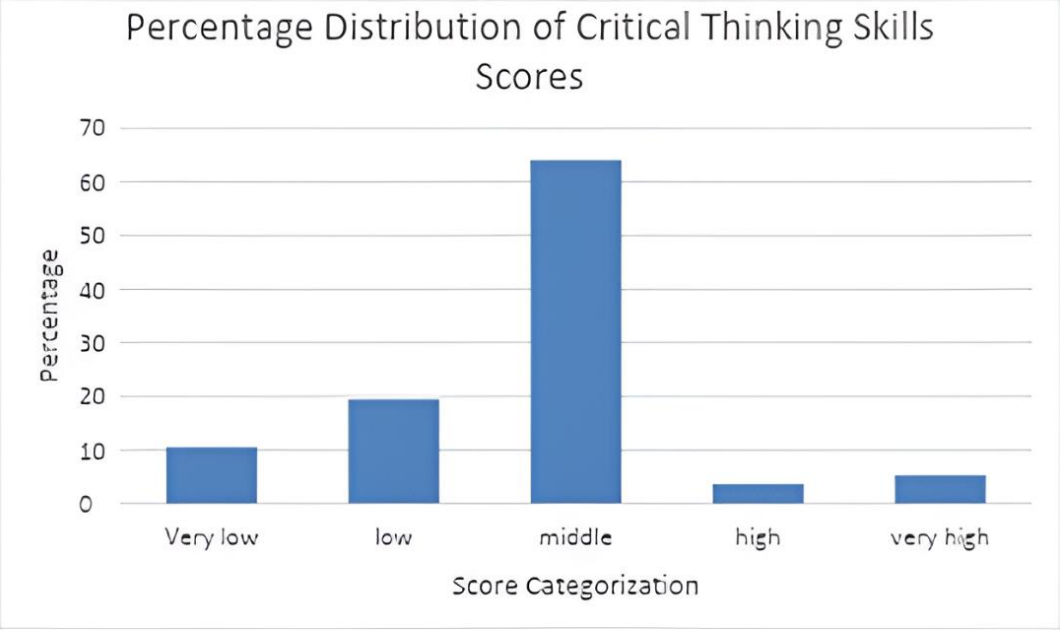

Fig.1. Percentage Distribution of Critical Thinking Skills Scores on Ecological Issues.

Source: Statistical calculations from primary data.

Table 5. Results of Data Analysis to Measure Comparison of MIN.

Group	Kruskal Wallis Test P-Value	Dunn Test	(hal.adj)	Z
MIN	2.2e-16 (There is a significant difference)	A vs B	5.908941e-24 (There is a significant difference)	10.2677423
		A vs C	5.656930e-09 (There is a significant difference)	6.1188008
		B vs C	9.526983e-03 (There is a significant difference)	-3.1581331
		A vs D	1.000000e+00 (There is not significantly different)	0,2743991
		B vs D	2.447004e-26 (There is a significant difference)	-10.7844141
		C vs D	2.023228e-09 (There is a significant difference)	-6.2806296

Source: Statistical calculations from primary data
Notes:
A: MIN Bogor Regency, West Java (digital format test)
B: MIN Palembang City, South Sumatra (paper format test)
C: MIN Sinjai Regency, South Sulawesi (paper format test)
D: MIN Balikpapan City, East Kalimantan (digital format test)

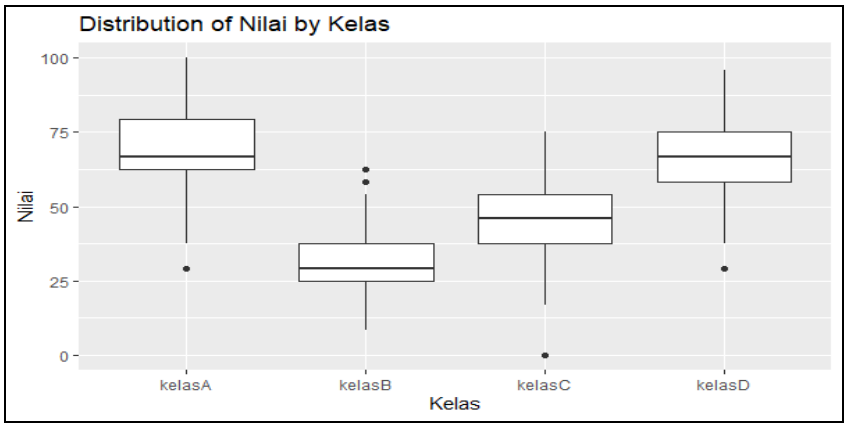

Fig. 2. Boxplot comparison between MIN.

Source: Statistical calculations from primary data.

Notes:
class A: MIN Bogor Regency, West Java (digital format test)
class B: MIN Palembang City, South Sumatra (paper format test)
class C: MIN Sinjai Regency, South Sulawesi (paper format test)
classD: MIN Balikpapan City, East Kalimantan (digital format test)

The results presented in Table 5 are reinforced by Figure 4, which shows that the critical thinking abilities of MIN Bogor Regency students are not significantly different from those of MIN Balikpapan students. In Table 4, you can see the P.adj value and Z value. The greater the P.adj value, the smaller than 0.05, the more significant the differences between groups, while the Z value shows the distance between groups. The average value of MIN for Bogor Regency is much higher than MIN for Sinjai Regency and MIN for Palembang City. Likewise, the average score for MIN

Balikpapan is much higher than the average critical thinking ability score for MIN Sinjai Regency and MIN Palembang City. The MIN of Sinjai Regency shows a slightly higher average value than the MIN of Palembang City.

3.3. Discussion

Seeing the results above that the average critical thinking skills of students are in the low to medium range and the results of the assessment of students' critical thinking skills are higher in digital format compared to paper format, this shows that Madrasahs need to pay attention to the presence of digital elements in the assessment process learning and improving the learning process that focuses on improving critical thinking skills learners.

Based on the data in Table 5 and Figure 2, it is known that the MIN values for Bogor Regency and MIN Balikpapan City, which use digital format, have significantly higher values compared to MIN Palembang and MIN Sinjai, which use paper format. The average value of MIN for Bogor Regency is much higher than the average value for MIN for Sinjai Regency and MIN for Palembang City. Likewise, the average score for MIN Balikpapan is much higher than the average critical thinking ability score for MIN Sinjai Regency and MIN Palembang City. The MIN of Sinjai Regency shows a slightly higher average value than the MIN of Palembang City. MIN Sinjai and MIN Bogor are in the district, while MIN Palembang and MIN Balikpapan are in the capital city. MIN Bogor Regency is close to the capital city of the Republic of Indonesia, DKI Jakarta, so the area's characteristics are suburban. It is known that the critical thinking abilities of students who live in urban areas are higher than the critical thinking abilities of students who live in rural areas (Kurniawan et al., 2023; Tanti et al., 2020). However, the results of MIN Palembang students' critical thinking ability profile are relatively low for urban MIN students. Critical thinking skills of elementary school students in Indonesia tend to be low due to student, teacher, and curriculum factors. Students often make mistakes in answering, rely on memory rather than deep understanding, and have low mathematical dispositions, such as a lack of curiosity (Arifuddin et al., 2023; Sarwanto et al., 2021)On the other hand, teachers tend to use less effective lecture learning models, have limited pedagogical knowledge, and face limited time and resources in integrating critical thinking skills (Sarwanto et al., 2021; Trisnani et al., 2024). Although the Merdeka Curriculum and literacy programs aim to improve these skills, their implementation is still not optimal (Jumanto et al., 2024). To overcome this, innovation is needed in learning models such as discovery learning, the use of mobile applications, as well as the development of a curriculum that better supports critical thinking skills (Isrokatun et al., 2022; Jumanto et al., 2024).

Factors that influence success in implementing written assessments are the readability of the questions and students' psychological factors apart from mastery of the material (Brüggemann et al., 2023a; Eno, 2011; Herling, 2009; Kurniawati & Istiyono, 2022; Piazzalunga et al., 2023; Sabil et al., 2024). In this research, the readability factor of the questions has been minimized through expert validation and reliability testing. The study provided shows that on multiple-choice tests, students who use Computer-Based Testing (CBT) with personal devices show improved performance compared to Paper-Based Testing (PBT) (Nardi & Ranieri, 2019). It may be due to several factors, including a positive relationship between students' self-efficacy with computers and their propensity to take electronic tests, as familiarity with using their own devices may increase this self-efficacy. Additionally, immediate feedback, a key benefit of CBT, supports better performance and reduces test anxiety. CBT also offers increased flexibility and control, allowing students to change answers right up to the end of the test, further reducing anxiety. These findings suggest that well-designed CBT systems, especially when using familiar personal devices, can improve performance by creating a more comfortable and less stressful testing environment ((Nardi & Ranieri, 2019), in line with the opinion of Clariana & Wallace (2002), who compared test results via CBT vs PBT. This study showed that higher achieving students benefited most from CBT and that content familiarity was related to test mode effects. Familiarity with computers is key in test mode effects, especially for unfamiliar content or low-achieving students. He argues that CBT can improve performance due to increased focus, the ability to review and change answers, and the alignment of test mode with instruction and that high-achieving students benefit most from this. Format, especially when they are familiar with computers.

From psychological factors, these findings align with research regarding the differences between working on paper-based and computer-based questions. It is known that working on paper-based questions can give students anxiety because they observe the number of pages that must be completed compared to digital-based assessments, which only focus on completing each item (Clariana & Smith, 1988; Eno, 2011). However, the results of other research by Brüggemann et al. (2023a) show no difference in cognitive load between students who use paper and digital format tests. In addition, students currently tend to interact longer with gadgets than with paper. This familiarity factor will make students more comfortable using digital media compared to paper media (Eno, 2011; Farber, 2017; McDonald & Fotakopoulou, 2023; Støle et al., 2020).

Experience in constructivist learning can improve critical thinking skills (Ashari & Sun, 2024; Guo et al., 2024; Yousefizadeh, 2023). In the learning process, various forms of evaluation are needed to measure the success of the learning process, as well as direct feedback and reflection so that learning runs effectively (Baviskar et al., 2009; Jonassen, 2013). Teachers who are skilled at using digital-based assessments will gain benefits and ease in implementing constructive learning. Applications based on gamification, such as Kahoot and Quiziss, can increase students' involvement and motivation in learning (Göksün & Gürsoy, 2019; Maraza-Quispe, 2024). In addition, the digital assessment allows a more adaptive approach (*adaptive assessment*), where the difficulty level of questions can be adjusted to students' abilities in real time (Owan et al., 2023; Swiecki et al., 2022). *Adaptive assessment* can increase the effectiveness of measuring critical thinking skills by providing challenges appropriate to students' cognitive capacities. In the madrasa context, this approach can be used to ensure that students with varying levels of understanding still receive appropriate challenges in developing their analytical and evaluative thinking skills. Thus, for students critical thinking skills to improve, facilities and infrastructure are needed to carry out assessments effectively *in real-time* and support for increasing teacher competency in utilizing digital assessments. The use of devices in strictly restricted school environments needs to be considered at certain times because it turns out that the use of devices can have a positive impact on supporting constructivist learning (Halkiopoulos & Gkintoni, 2024).

What has not been measured from this research is whether the results of critical thinking skills with multiple responses or two-tier multiple choice in digital format can be better than in paper format (Fajari & others, 2021; Ku, 2009; Rintayati et al., 2021). This format is more effective for measuring critical thinking skills than multiple choice (Ku, 2009). The longer the duration of the test, the more words students must read, and the need for students to think longer are likely toroduce different test results. Research shows that the reading ability test results for 10-year-old students using paper produce better understanding than those using tablets, especially for female students (Arenas & Palaoag, 2018). Similar results were obtained by Carpenter and Alloway (2019), who showed that paper-format tests show improved working memory performance.

In contrast, computer-based tests may produce lower scores due to developmental differences, increased cognitive workload, test mode effects, and access. Individual towards technology. Lee et al. (2024) examined brain activity in children who used tablet media compared to paper, explaining that prefrontal cortex activity during the encoding phase was higher in digital media users than in paper media users to access knowledge. However, energy efficiency in brain use is higher when using paper mode compared to digital mode to achieve the same goal. It is in line with research by Brüggemann et al. (2023a), which found that initial motivation for taking tests in digital format decreased as the test was taken. Thus, further studies are needed to analyse clear boundaries regarding using digital devices for critical thinking skills assessment needs in schools.

Based on the results of this study, it is suggested that the use of devices can be applied to Madrasahs for the implementation of learning but with the right strategy according to the learning objectives. The use of devices is needed so that students can explore it as interesting learning media and can explore high-level thinking skills. However, the use of these devices needs to be studied further to determine the limits of their use. For example, is the use of gadgets effective in reading comprehension compared to paper-based ones? Paper media or devices that better support cognitive and motor development also need to be considered so that the use of devices in learning does not inhibit thinking skills but supports the learning process.

4. Conclusion

Based on the results of this research, it is known that digital-based assessment formats produce higher critical thinking skills scores compared to paper-based assessments at Madrasah Ibtidaiyah in Indonesia. This difference shows that test format factors can influence the results of critical thinking skills assessments, which can be caused by students' psychological factors and familiarity with technology. However, this research still has limitations in measuring other factors that can influence differences in results, such as students' experience with digital technology in daily learning and the readiness of madrasah infrastructure to implement digital assessments.

Therefore, further research is needed to 1) Analyze the impact of using multiple test formats on other forms of instruments to measure critical thinking skills, 2) Examine solutions and strategies for using digital assessments to improve students' critical thinking skills within limited facilities and infrastructure as well as limited access for Madrasah teachers receive assistive and ongoing professional development training.

This research suggests that educational policies in madrasas must consider using technology in assessment to increase the effectiveness of critical thinking-based learning. In addition, teachers need training to integrate digital assessments into more comprehensive learning strategies. Thus, this study provides important insights into the effectiveness of digital assessment in Islamic education and can be the basis for developing policies to improve the quality of learning in madrasas on an ongoing basis.

Acknowledgements

This research is supported by the Ministry of Religious Affairs for Raden Fatah Palembang State Islamic University through a research grant scheme for 2024, as stated in Decree Number 779 of 2024.

References

Almeida, B., Santos, M., & Justi, R. (2023). Aspects and abilities of science literacy in the context of nature of science teaching. *Science \& Education*, *32*(3), 567–587.

Anggraeni, D. M., Prahani, B. K., Suprapto, N., Shofiyah, N., & Jatmiko, B. (2023). Systematic review of problem based learning research in fostering critical thinking skills. *Thinking Skills and Creativity*, *49*, 101334. https://doi.org/https://doi.org/10.1016/j.tsc.2023.101334

Arenas, M. R., & Palaoag, T. D. (2018). Gamified learning system: A motivation for exceptional learners of SPED-baguio for better engagement. *Journal of Advanced Research in Dynamical and Control Systems*, *10*(11), 216 – 222. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85074669686&partnerID=40&md5=ec4f0e976f1bef2043b4ca194a43a4f2

Arifuddin, A., Wahyudin, W., & Prabawanto, S. (2023). THE EFFECT OF STUDENTS' MATHEMATICAL DISPOSITION ON MATHEMATICAL CRITICAL THINKING ABILITY OF ELEMENTARY SCHOOL STUDENTS. *Lentera Pendidikan : Jurnal Ilmu Tarbiyah Dan Keguruan*. https://doi.org/10.24252/lp.2023v26n2i4

Ashari, Z. M., & Sun, D. (2024). Project Based Learning on Promoting Children's Critical Thinking Skills: A Systematic Review. *International Journal of Academic Research in Progressive Education and Development*. https://doi.org/10.6007/ijarped/v13-i3/21732

Azra, A. (2021). Understanding Indonesia's third way Islam. In *Wars of Ideas: Theology, Interpretation and Power in the Muslim World* (p. 73). Rowman \& Littlefield.

Bacak, J., Martin, F., Ahlgrim-Delzell, L., Polly, D., & Wang, W. (2022). Elementary Educator Perceptions of Student Digital Safety Based on Technology Use in the Classroom. *Computers in the Schools*, *39*, 186–202. https://doi.org/10.1080/07380569.2022.2071233

Bainiashova, S. (2023). Content analysis of the concept of" information literacy". *Zhytomyr Ivan Franko State University Journal. Pedagogical Sciences*, *3 (114)*, 5–15.

Baviskar, S. N., Hartle, R., & Whitney, T. (2009). Essential Criteria to Characterize Constructivist Teaching: Derived from a review of the literature and applied to five constructivist-teaching method articles. *International Journal of Science Education*, *31*, 541–550. https://doi.org/10.1080/09500690701731121

Brüggemann, T., Ludewig, U., Lorenz, R., & McElvany, N. (2023a). Effects of mode and medium in reading comprehension tests on cognitive load. *Computers \& Education*, *192*, 104649.

Brüggemann, T., Ludewig, U., Lorenz, R., & McElvany, N. (2023b). Effects of test mode and medium on elementary school students' test experience. *European Journal of Psychological Assessment*.

Carpenter, R., & Alloway, T. (2019). Computer Versus Paper-Based Testing: Are They Equivalent When it Comes to Working Memory? *Journal of Psychoeducational Assessment*, *37*, 382–394. https://doi.org/10.1177/0734282918761496

Christensen, E. W., & Bailey, J. R. (1997). A source accessibility effect on media selection. *Management Communication Quarterly*, *10*(3), 373–387.

Ciampa, K., & Gallagher, T. (2013). Getting in Touch: Use of Mobile Devices in the Elementary Classroom. *Computers in the Schools*, *30*, 309–328. https://doi.org/10.1080/07380569.2013.846716

Clariana, R. B., & Smith, L. (1988). Learning Style Shifts in Computer-Assisted Instructional Settings. *The Annual Meeting of the International Association for Computers in Education Held in Conjunction with the American Educational Research Association*, 20.

Dewi, N. K. L., Utami, I. G. A. L. P., & Santosa, M. H. (2022). Development of Digital Supplementary Material Using Quizizz-Based Learning Media in Intensive English Course. *Jurnal Inovasi Dan Teknologi Pembelajaran (JINOTEP): Kajian Dan Riset Dalam Teknologi Pembelajaran*, *9*(1), 14–22.

Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, *18*(3), 4–10.

Eno, L. P. (2011). *Comparing the Reading Performance of High-Achieving Adolescents: Computer-Based Testing versus Paper/Pencil.* Seton Hall University.

Erwin, T. D. (2000). *The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing*. https://nces.ed.gov/pubs2000/2000195.pdf#page=20.11

Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report)*.

Facione, P. A., & others. (2011). Critical thinking: What it is and why it counts. *Insight Assessment*, *1*(1), 1–23.

Fajari, L. E. W., & others. (2021). Open-Ended Questions to Assess Critical-Thinking Skills in Indonesian Elementary School. *International Journal of Instruction*, *14*(1), 615–630.

Farber, M. (2017). *Gamify Your Classroom: A Field Guide to {Game-Based} Learning (New Literacies and Digital Epistemologies)*.

Ferrary, C. H., Kawuryan, S. P., & others. (2023). How are Critical Thinking Skills and Integration of IPS Problems in Elementary Schools? *Mimbar Ilmu*, *28*(1), 147–155.

Fitanti, S. Y., Prahani, B. K., & Suryanti, S. (2024). Profile of Critical Thinking Skills of Grade IV Elementary School Students on Energy Source Material. *Scaffolding: Jurnal Pendidikan Islam Dan Multikulturalisme*, *6*(3), 58–74.

Gelerstein, D., Del Rio, R., Nussbaum, M., Chiuminatto, P., & López, X. (2016). Designing and implementing a test for measuring critical thinking in primary school. *Thinking Skills and Creativity*, *20*, 40–49.

Gendenjamts, S. (2023). MEASURING HIGHER-ORDER THINKING SKILLS IN SCIENCE AMONG PRIMARY SCHOOL STUDENTS USING ITEM RESPONSE THEORY. *European Journal of Education Studies*, *10*(12).

Göksün, D. O., & Gürsoy, G. (2019). Comparing success and engagement in gamified learning experiences via Kahoot and Quizizz. *Computers \& Education*, *135*, 15–29.

Guo, R., Jantharajit, N., & Thongpanit, P. (2024). Construct an instructional approach based on collaborative learning and reflective learning for enhance students' analytical thinking and critical thinking skills. *Asian Journal of Contemporary Education*. https://doi.org/10.55493/5052.v8i2.5184

Halkiopoulos, C., & Gkintoni, E. (2024). Leveraging AI in e-learning: Personalized learning and adaptive assessment through cognitive neuropsychology—A systematic analysis. *Electronics*, *13*(18), 3762.

Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology press.

Hasanah, S. N., Sunarno, W., & Prayitno, B. A. (2021). Improving students' critical thinking skills trough contextual teaching and learning science module. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *6*(2), 106.

Herling, C. S. (2009). *No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分 析 Title*. *369*(1), 1689–1699. http://dx.doi.org/10.1016/j.jsames.2011.03.003%0Ahttps://doi.org/10.1016/j.gr.2017.08.001%0Ahttp://dx.doi.org/10.1016/j.precamres.2014.12.018%0Ahttp://dx.doi.org/10.1016/j.precamres.2011.08.005%0Ahttp://dx.doi.org/10.1080/00206814.2014.902757%0Ahttp://dx.

Hitchcock, D. (2018). *Critical thinking*. Stanford Encyclopedia of Philosophy Archive. https://stanford.library.sydney.edu.au/archives/spr2019/entries/critical-thinking/

Hollett, E. C. (2022). *Analytical Thinking for Advances Learners, Grades 3-5*. ROUTLEDGE.

Hsu, F.-H., Lin, I.-H., Yeh, H.-C., & Chen, N.-S. (2022). Effect of Socratic Reflection Prompts via video-based learning system on elementary school students' critical thinking skills. *Computers & Education*, *183*, 104497. https://doi.org/https://doi.org/10.1016/j.compedu.2022.104497

Hussein, M. H., Ow, S. H., Cheong, L. S., & Thong, M. (2019). A Digital Game-Based Learning Method to Improve Students' Critical Thinking Skills in Elementary Science. *IEEE Access*, *7*, 96309–96318. https://doi.org/10.1109/ACCESS.2019.2929089

Ishartono, N., Desstya, A., Prayitno, H., & Sidiq, Y. (2021). The Quality of HOTS-Based Science Questions Developed by Indonesian Elementary School Teachers. *Journal of Education Technology*. https://doi.org/10.23887/jet.v5i2.33813

Isrokatun, I., Hanifah, N., Rohman, Y. A., Rosmiati, R., & Khoerunnisah, R. (2022). Enhancing critical thinking skills for low-grade elementary school students using mobile apps. *Cypriot Journal of Educational Sciences*, *17*(9), 3218 – 3237. https://doi.org/10.18844/cjes.v17i9.7422

Jeong, S., Sherman, B., & Tippins, D. J. (2021). The Anthropocene as we know it: Posthumanism, science education and scientific literacy as a path to sustainability. *Cultural Studies of Science Education*, *16*, 805–820.

Jonassen, D. H. (2013). Evaluating constructivistic learning. In *Constructivism and the technology of instruction* (pp. 137–148). Routledge.

Juliyantika, T., & Batubara, H. H. (2022). Tren Penelitian Keterampilan Berpikir Kritis pada Jurnal Pendidikan Dasar di Indonesia. *Jurnal Basicedu*, *6*(3), 4731–4744.

Jumanto, Sa'ud, U. S., & Sopandi, W. (2024). Profile of Critical Thinking Skills of Elementary School Students in Surakarta City Based on Elements Curriculum Merdeka. *SHS Web of Conferences*. https://doi.org/10.1051/shsconf/202418201007

Kannengiesser, U., & Gero, J. S. (2019). Design thinking, fast and slow: a framework for Kahneman's dual-system theory in design. *Design Science*, *5*.

Kawuryan, S. P., Sayuti, S., & Aman, A. (2022). A DESCRIPTIVE STUDY OF CRITICAL THINKING ABILITIES OF ELEMENTARY SCHOOL STUDENTS. *Jurnal Cakrawala Pendidikan*. https://doi.org/10.21831/cp.v41i1.44322

Kennedy, M., Fisher, M. B., & Ennis, R. H. (1991). Critical thinking: Literature review and needed research. *Educational Values and Cognitive Instruction: Implications for Reform*, *2*, 11–40.

Kilickaya, F., & Ersoy, M. A. (2016). Teaching how to write instructional objectives to pre-service language teachers through the ABCD Model. *New Insights into Language Teaching and Learning Practices*, 41–52.

Kisilowska-Szurmińska, M. (2023). Information Overload as a Burden and a Challenge. What Can We Learn for Information Literacy? *European Conference on Information Literacy*, 161–172.

Ku, K. Y. (2009). Assessing Students' Critical Thinking Performance: Urging for Measurements Using Multi-Response Format. *Thinking Skills and Creativity*, *4*, 70–76. https://doi.org/10.1016/J.TSC.2009.02.001

Kurniawan, D. A., Wirayuda, R. P., Putri, W. A., Rini, E. F. S., Ginting, A. A. B., Ratnawati, T., & others. (2023). Impact of science process skills on thinking skills in rural and urban schools. *International Journal of Instruction*, *16*(2), 803–822.

Kurniawati, & Istiyono, E. (2022). Developing a Critical Thinking Test in Social Science for Grade V Elementary School. *Advances in Social Science, Education and Humanities Research*. https://doi.org/10.2991/assehr.k.220129.040

Kusnanto, N., Sukristyanto, A., & Rochim, A. I. (2023). Relevance Of National Education Policies As An Effort To Improve The Quality Of Madrasah Tsanawiyah Education Services. *THE SPIRIT OF SOCIETY JOURNAL: International Journal of Society Development and Engagement*, *6*(2), 136–151.

Lee, S.-A., Hong, J.-H., Kim, N.-Y., Min, H.-M., Yang, H.-M., Lee, S.-H., Choi, S.-J., & Park, J.-H. (2024). Unveiling Neurocognitive Disparities in Encoding and Retrieval between Paper and Digital Tablet-Based Learning. *Brain Sciences*, *14*(1), 76.

Lin, I.-H., Chew, S., & Chen, N. (2017). A Vocal Assessment Approach to Measure Elementary School Students' Critical Thinking Skills. *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, 419–421. https://doi.org/10.1109/ICALT.2017.30

Manurung, A. S., & Pappachan, P. (2025). The role of discovery learning in efforts to develop students' critical thinking abilities. *Journal of Education and Learning*, *19*(1), 46 – 53. https://doi.org/10.11591/edulearn.v19i1.21788

Maraza-Quispe, B. (2024). Impact of the Use of Gamified Online Tools: A Study with Kahoot and Quizizz in the Educational Context. *International Journal of Information and Education Technology*. https://doi.org/10.18178/ijiet.2024.14.1.2033

Mariana, D., & Helmi, A. M. (2022). Madrasah Sebagai Lembaga Pendidikan di Indonesia. *Jurnal Pendidikan Tambusai*, *6*(1), 1907–1919.

Martani, E. (2023). Improving the quality of madrasah education through madrasah resource management. *Dawuh Guru: Jurnal Pendidikan MI/SD*, *3*(1), 25–38.

Mason, M. (2007). Critical thinking and learning. *Educational Philosophy and Theory*, *39*(4), 339–349.

Mastuti, A. G., Abdillah, A., Sehuwaky, N., & Risahondua, R. (2022). Revealing students' critical thinking ability according to facione's theory. *Al-Jabar: Jurnal Pendidikan Matematika*, *13*(2), 261–272.

McDonald, S., & Fotakopoulou, O. (2023). Narrowing the Digital Divide in Early Maths: How Different Modes of Assessment Influence Young Children's Mathematical Test Scores. *Early Education and Development*, *35*, 132–149. https://doi.org/10.1080/10409289.2023.2260683

Melawati, Y., Rosidin, U., Nurhanurawati, N., & Handoko, H. (2023). Development of A Higher-Order Thinking Skills (HOTS) Instrument to Measure Student's Critical Thinking Abilities in Class V Elementary School. *International Journal of Current Science Research and Review*. https://doi.org/10.47191/ijcsrr/v6-i12-70

Millah, M. F., Jaenudin, R., Nurhayati, E., Mulyani, S., Nurmaulidah, E. F., & Nasrudin, D. (2022). AKMI: Initial Knowledge, misconceptions, and readiness of madrasa teachers. *International Conference on Madrasah Reform 2021 (ICMR 2021)*, 154–159.

Nardi, A., & Ranieri, M. (2019). Comparing paper-based and electronic multiple-choice examinations with personal devices: Impact on students' performance, self-efficacy and satisfaction. *British Journal of Educational Technology*, *50*(3), 1495–1506.

Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(8), em2307.

Park, J., & Niyozov, S. (2008). Madrasa education in South Asia and Southeast Asia: current issues and debates. *Asia Pacific Journal of Education*, *28*(4), 323–351.

Perry, K., Meissel, K., & Hill, M. F. (2022). Rebooting assessment. Exploring the challenges and benefits of shifting from pen-and-paper to computer in summative assessment. *Educational Research Review*, *36*, 100451. https://doi.org/https://doi.org/10.1016/j.edurev.2022.100451

Piazzalunga, C., Donati, A., & Ferrante, S. (2023). Co-design of a points-based reward system to boost motivation in children and improve adherence to learning serious games. *Convegno Nazionale Di Bioingegneria*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85175799565&partnerID=40&md5=88c03a9a748887751e696c27735c19fa

Primasari, R., Miarsyah, M., & Rusdi, R. (2020). Science literacy, critical thinking skill, and motivation: A correlational study. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, *6*(2), 273–282.

Queiruga-Dios, M. Á., López-Iñesta, E., Diez-Ojeda, M., Sáiz-Manzanares, M. C., & Vázquez Dorr\'\io, J. B. (2020). Citizen science for scientific literacy and the attainment of sustainable development goals in formal education. *Sustainability*, *12*(10), 4283.

Ridzal, D. A., & Haswan, H. (2023). Analysis of the correlation between science literacy and critical thinking of grade eight students in the circulatory system. *Jurnal Pijar Mipa*, *18*(1), 1–5.

Rintayati, P., Lukitasari, H., & Syawaludin, A. (2021). Development of Two-Tier Multiple Choice Test to Assess Indonesian Elementary Students' Higher-Order Thinking Skills. *International Journal of Instruction*. https://doi.org/10.29333/iji.2021.14133a

Rosa, N. (2022). Selain ANBK, Madrasah juga Wajib Ikut AKMI, Apa Itu? *Detikedu*. https://www.detik.com/edu/sekolah/d-6289196/selain-anbk-madrasah-juga-wajib-ikut-akmi-apa-itu

Sabil, H., Asrial, Syahrial, Kurniawan, D. A., Perdana, R., Nawahdani, A. M., Widodi, B., & Belajar, P. M. (2024). Students' Learning Interest Using Web-Based Assessment at The Elementary School Level. *Jurnal Ilmiah Sekolah Dasar*. https://doi.org/10.23887/jisd.v7i4.59737

Sarwanto, S., Fajari, L. E. W., & Chumdari, C. (2021). Critical thinking skills and their impacts on elementary school students. *Malaysian Journal of Learning and Instruction*, *18*(2), 161. https://doi.org/10.32890/mjli2021.18.2.6

Shin-Jia, H., & Tien-Chi, H. (2024). Integrating the United Nations' sustainable development goals into STEAM education: A problem-based learning approach in elementary and secondary education. *Journal of Textbook Research*, *17*(2), 79.

Shute, V., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *J. Comput. Assist. Learn.*, *33*, 1–19. https://doi.org/10.1111/jcal.12172

Shyr, W., Hsieh, Y., & Chen, C.-H. (2021). The Effects of Peer-Based Instant Response System to Promote Learning Performance, Intrinsic Motivation and Self-Efficacy. *Sustainability*. https://doi.org/10.3390/SU13084320

Sofanudin, A., Rokhman, F., & others. (2016). Quality-Oriented Management of Educational Innovation at Madrasah Ibtidaiyah. *Journal of Education and Practice*, *7*(27), 176–180.

Støle, H., Mangen, A., & Schwippert, K. (2020). Assessing children's reading comprehension on paper and screen: A mode-effect study. *Comput. Educ.*, *151*, 103861. https://doi.org/10.1016/j.compedu.2020.103861

Sumarni, W., & Kadarwati, S. (2020). Ethno-stem project-based learning: Its impact to critical and creative thinking skills. *Jurnal Pendidikan IPA Indonesia*, *9*(1), 11–21.

Sutiani, A., Situmorang, M., & Silalahi, A. (2021). Implementation of an inquiry learning model with science literacy to improve student critical thinking skills. *International Journal of Instruction*, *14*(2), 117–138.

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*, 100075.

Tanti, T., Kurniawan, D. A., Kuswanto, K., Utami, W., & Wardhana, I. (2020). Science process skills and critical thinking in science: Urban and rural disparity. *Jurnal Pendidikan IPA Indonesia*, *9*(4), 489–498.

Trisnani, N., Retnawati, H., & Wuryandani, W. (2024). Challenges of Indonesian elementary school mathematics teachers in integrating critical thinking into the classroom. *Journal on Mathematics Education*. https://doi.org/10.22342/jme.v15i3.pp905-924

Ubaidillah, M., Hartono, H., Marwoto, P., Wiyanto, W., & Subali, B. (2023). How to Improve Critical Thinking in Physics Learning? A Systematic Literature Review. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, *28*, 161–187.

Utaminingsih, S., Amalia, I., & Sumaji, S. (2024). Management of Mathematics Learning Based on Interactive Digital Worksheets to Improve Students' Critical Thinking Ability. *Journal of Curriculum and Teaching*. https://doi.org/10.5430/jct.v13n1p159

Vieira, R., & Tenreiro-Vieira, C. (2016). Fostering Scientific Literacy and Critical Thinking in Elementary Science Education. *International Journal of Science and Mathematics Education*, *14*, 659–680. https://doi.org/10.1007/S10763-014-9605-2

Wani, S. A., & Hussian, Z. (2024). Developing Critical Thinking Skills: Encouraging Analytical and Creative Thinking. In *Transforming Education for Personalized Learning* (pp. 114–130). IGI Global.

Widhiarso, W., & Ridho, A. (2022). Validation of Setting and Design of Multi-Stage Testing (MST) to Portray Students' Achievement on Reading Literacy in AKMI 2021. *International Conference on Madrasah Reform 2021 (ICMR 2021)*, 6–10.

WorldBank. (2024). *Realizing Education's Promise: Support to Indonesia's Ministry of Religious Affairs for Improved Quality of Education*. https://projects.worldbank.org/en/projects-operations/project-detail/P168076

Xu, L., Abdullah, T., Liu, Q., & Ain Mohd Shahroom, N. (2023). *Effects of Critical Thinking Disposition on Foreign Language Proficiency in Foreign Language Learning: Evidence from China*.

Yousefizadeh, S. (2023). Evaluation of Curriculum Programs Based on Constructivist Learning Theory and the Development of Critical Thinking. *Journal of Study and Innovation in Education and Development*. https://doi.org/10.61838/jsied.3.1.4

Zarkasy, A., Chotimah, C., & Mahzuna, L. (2023). IMPLEMENTATION OF MASSIVE ONLINE OPEN COURSE (MOOC) THROUGH PINTAR IN IMPROVING TEACHER COMPETENCE. *Annual International Conference on Islamic Education for Students*, *2*(1), 665–674.