

Comparing Paper-Based and Computer-Based Testing: A Systematic Literature Review on Their Effectiveness in Performance Assessment

Meili Yanti¹, Irmawati M², Dea Santika Rahayu³, Niken Yuni Astiti⁴

¹ Science Education, Universitas Sulawesi Barat, Indonesia

² Primary Teacher Education, Universitas Sulawesi Barat, Indonesia

³ Science Education, Universitas Tidar, Indonesia

⁴ Primary Teacher Education, Universitas Lampung, Indonesia

¹meiliyanti@unsulbar.ac.id, ²irmawati.m@unsulbar.ac.id, ³dea.santika@untidar.ac.id,

⁴nikenastiti@fkip.unila.ac.id

ARTICLE INFO

Article History:

Received : 30-04-2026

Revised : 29-08-2026

Accepted : 13-09-2026

Online : 15-10-2026

Keywords:

Evaluation;

CBT;

Motivation;

PBT;

Performance Assesment.



ABSTRACT

The rapid advancement of technology has significantly influenced various aspects of human life. While these developments offer substantial convenience, they also present certain challenges, particularly during the early stages of adoption. In the field of education, technology integration is not only evident in instructional processes but also in assessment practices. One prominent example is the implementation of Computer-Based Tests (CBT), which provide several advantages; however, their effectiveness in evaluating student performance requires further examination. Accordingly, this study aims to compare the traditional Paper-Based Test (PBT) with CBT in terms of their effectiveness in performance assessment within the learning process. This study employs a Systematic Literature Review (SLR) approach guided by two research questions: (P1) How does the effectiveness of Paper-Based Test (PBT) compare with Computer-Based Test (CBT)? and (P2) How do student performances differ between PBT and CBT? The findings from the SLR indicate mixed results. Several studies reveal that students tend to show a higher preference for CBT over PBT, accompanied by increased motivation. Nevertheless, both assessment methods appear to have no significant impact on students' performance outcomes. Additional findings and analyses are further elaborated in this study.



Crossref



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

A. INTRODUCTION

Learning is a process aimed at achieving change, particularly changes that are relatively permanent (Munna & Kalam, 2021). These changes are reflected in learning objectives, which encompass the acquisition of knowledge, attitudes, and values that students are expected to apply in their future lives (Budiastuti et al., 2021). Therefore, learning is continuously designed to align with students' needs, which are also closely related to workforce demands.

In the era of technological advancement, the expectations for graduates have significantly increased. These expectations are no longer limited to knowledge acquisition but extend to the development of specific skills (García-Pérez et al., 2021). Consequently, learning processes must be carefully designed to meet these evolving demands.

Assessment plays a critical role in determining whether these competencies have been successfully achieved. High-quality assessment practices enable educators to generate

valid and reliable conclusions regarding student performance and learning outcomes (Hassler Hallstedt & Ghaderi, 2018).

The integration of computers and internet technologies has fundamentally transformed educational practices, particularly in pedagogy, curriculum development, and assessment strategies (Brown, 2019; Carpenter & Alloway, 2019). One of the most significant developments in assessment is the shift from traditional Paper-Based Tests (PBT) to Computer-Based Tests (CBT) (Hakim, 2018). Digital developments enable teachers at all levels of education to initiate technology-based assessment systems (Yanti et al., 2019)

Furthermore, CBT can be used repeatedly and boosts student motivation, thereby increasing students' willingness to take tests and their level of engagement (Piaw, 2012). In addition, another advantage of CBT is its ability to reduce exam time and foster intrinsic and social motivation (Jalali et al., 2014). The appealing interfaces of computers and smartphones make them promising platforms for technology-based assessment.

CBT offers several advantages over traditional assessment methods. It provides greater consistency in administration, reduces human error in scoring, and allows for immediate feedback. Moreover, CBT environments have been associated with increased student motivation and engagement, as they incorporate interactive and user-friendly interfaces (Hosseini et al., 2014). Some schools have established CBT as an assessment method that utilizes computers in every stage of the exam: preparation, administration, scoring, and displaying student scores (Awofala, 2020).

Despite some drawbacks, computer-based testing is gaining popularity due to the benefits it offers, such as effective administration, flexible scheduling, and immediate feedback (Khoshsima & Toroujeni, 2017). Computer-Based Testing (CBT) has the capacity to assess a wide range of student competencies, including skills, knowledge, and conceptual understanding within a given subject area (Sobremisana & Aragon, 2016). Preparing students to engage with technology-based assessments is therefore essential, as it equips them with the competencies required to navigate an increasingly digitalized and computer-driven future (Ferdiansyah, 2016).

Previous studies can be mapped into four recurring themes. First, evidence regarding performance and score comparability is mixed. Several studies reported no statistically significant difference between CBT and PBT (Awofala, 2020; Bhoola-Patel & Laher, 2011; Karay et al., 2015; Khoshsima & Toroujeni, 2017), whereas other studies found higher performance in PBT, particularly for working-memory and language-related tasks (Carpenter & Alloway, 2019; Hosseini et al., 2014; Jalali et al., 2014). Conversely, some studies found advantages for digitally delivered assessment in particular contexts (Hakim, 2018; Sobremisana & Aragon, 2016; Yarahmadzahi & Goodarzi, 2020). These variations suggest that test mode effects may depend on the construct assessed, task format, participant characteristics, and implementation conditions.

Second, motivation and preference do not consistently correspond to performance. CBT has been associated with greater curiosity, involvement, comfort, and preference, yet increased motivation has not always produced higher scores (Piaw, 2012). Third, usability shapes the testing experience. Participants may value clearer navigation, easy answer revision, faster completion, and immediate results in CBT, while PBT allows annotation, familiar reading strategies, and independence from electricity or hardware (Jamiludin et al., 2017; Khoshsima & Toroujeni, 2017; Uukkivi et al., 2020).

Fourth, equity and fairness emerge when students differ in digital access, computer familiarity, age, or technological confidence. A testing interface may introduce construct-irrelevant cognitive load or anxiety, while inadequate infrastructure may expose some

participants to disruptions that are unrelated to the competence being assessed (Bhoola-Patel & Laher, 2011; Carpenter & Alloway, 2019). However, fairness is often treated as a contextual observation rather than a central analytical outcome.

The reviewed literature therefore remains fragmented across performance, motivation, usability, and equity. A synthesis is needed to clarify whether apparent score equivalence supports comparable interpretation and under which conditions CBT can be implemented fairly. This review addresses that need by examining score comparability and identifying motivational, usability, and access-related conditions that influence the validity and fairness of CBT and PBT.

B. METHODS

This study employed a systematic literature review to identify, appraise, and synthesize empirical studies that directly compared PBT and CBT. The review process followed the planning, conducting, and reporting principles described by Kitchenham et al. (2009), while reporting was structured using PRISMA 2020 (Page et al., 2021). The review addressed the following questions:

RQ1: To what extent do CBT and PBT produce comparable student performance outcomes?

RQ2: What motivational, usability, and equity-related factors influence the validity and fairness of CBT and PBT implementation?

By following this framework, the review ensures transparency, replicability, and reliability in analyzing the selected literature (Kitchenham et al., 2009). The stages of the review process consist of three main phases: planning, conducting, and reporting the results.

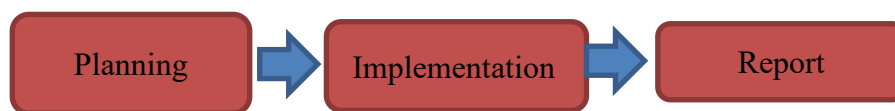


Figure 1. Research Stages

The systematic process illustrated in Figure 1 is further elaborated by Patricio Torres-Carrión, who provides a more detailed explanation of each stage within the review framework (Torres-Carrión et al., 2018).

The conducting stage involves the process of selecting relevant articles based on predefined keywords, inclusion and exclusion criteria, and a final screening procedure. Initially, the researchers determined a set of keywords to identify relevant studies. The keywords “paper-based test,” “computer-based test,” and “performance assessment” were used to search across major academic databases, including Google Scholar, ERIC (Education Resources Information Center), and Web of Science.

The initial search yielded a substantial number of studies, with approximately 4,120,000 articles identified in Google Scholar, 50,324 articles in ERIC, and 1,200 articles in Web of Science containing relevant information. These results indicate the extensive body of literature available on the topic.

Subsequently, the identified articles were further screened based on predefined inclusion and exclusion criteria, as presented in Table 1, to ensure the relevance and quality of the selected studies.

Table 1. *Inclusion dan Exclusion Criteria*

Inclusion Criteria	Exclusion Criteria
Using English	Using a language other than English
Published year 2010-2022	Published before 2010
It is within the scope of education	<i>Outside the scope of Education</i>
Empirical research	Not empirical research

Based on the selection process outlined in Table 1, a set of articles that met the inclusion criteria was obtained. Subsequently, a final screening was conducted to assess the relevance of these articles to the research topic. During this stage, the selected studies were categorized according to the research questions they addressed, namely those focusing on the first research question (P1) and those addressing the second research question (P2). The overall article selection process is illustrated in Figure 2.

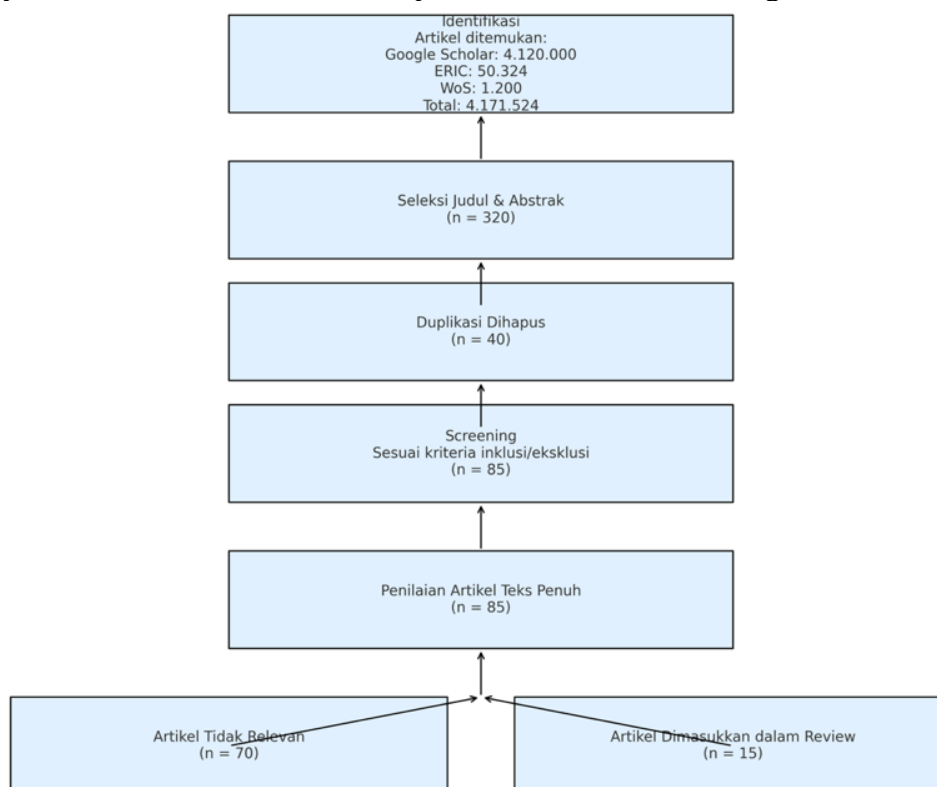


Figure 2. Article Selection Process

For each included study, the authors extracted the publication details, country and educational context, participants, sample size, research design, assessment modes, measured outcomes, and principal findings. Findings were coded deductively into four dimensions: (1) performance and score comparability, (2) motivation and preference, (3) usability and implementation efficiency, and (4) equity, accessibility, and fairness. A narrative synthesis was used because the studies differed in participants, instruments, subjects, and outcome measures.

C. RESULT AND DISCUSSION

1. Literature Review Result

Table 2. Thematic Analysis

Dimension	Evidence/advantage for CBT	Evidence/advantage for PBT	Integrated interpretation
Performance and score comparability	CBT advantage in selected contexts: Hakim (2018), Sobremisana & Aragon (2016), Yarahmadzahi & Goodarzi (2020).	PBT advantage for some language/working-memory tasks: Carpenter & Alloway (2019), Hosseini et al. (2014), Jalali et al. (2014).	No significant mode effect in several studies: Awofala (2020), Bhoola-Patel & Laher (2011), Karay et al. (2015), Khoshshima & Toroujeni (2017), Piaw (2012).
Motivation and preference	CBT often increased motivation, comfort, and preference: Piaw (2012), Khoshshima & Toroujeni (2017).	PBT preferred where learners lacked computer experience or valued familiar reading: Hosseini et al. (2014), Jamiludin et al. (2017).	Preference and motivation did not consistently predict performance.
Usability and implementation	Faster completion, answer revision, immediate results, automated scoring, and efficient storage.	Annotation, deep reading, familiar navigation, and independence from electricity, devices, and connectivity.	Usability depends on interface design, task type, training, and infrastructure.
Equity and fairness	Standardized delivery may support large-scale administration when access and training are adequate.	Limited access, digital unfamiliarity, developmental differences, screen fatigue, and technical disruption may disadvantage some learners.	Direct evidence on fairness remains limited; most studies treated equity as a contextual factor rather than a primary outcome.
The analysis of the research questions is based on the data systematically tabulated in the following table.			

Table 2. Analysis Result on the first research question (P1)

Title and authors	Results
Comparability of Test Results of Computer Based Tests (CBT) and Paper and Pencil Tests (PPT) among English Language Learners in Iran (Hosseini et al., 2014)	Students prefer the display of tests on the computer, which is about 53% compared to PPT. There were three students who stated that they preferred the PPT exam because they had no experience with CBT. The other two chose PPT because of an unpleasant experience, such as computer lessons, which they considered boring.
Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong? Chua Yan Piawa (2012) (Piaw, 2012)	Students show fatigue during the posttest using PBT so they get a low motivation score compared to when using CBT. The level of motivation has no correlation with student performance scores. This means that the selection of CBT and PBT does not make a difference in increasing the score.

Computer-based vs. Paper-based Writing Tests: A Comparative Study of Test Taking Strategies Masoud Zoghi (2013)	The most common strategies used by students in PBT and CBT are organization, planning, sense of reading, and evaluation. Meanwhile, a strategy that is rarely used is rehearsing.
Investigating the Role of Formative Mobile Based Assessment in Vocabulary Learning of Pre-Intermediate EFL Learners in Comparison with Paper Based Assessment (Yarahmadzahi & Goodarzi, 2020)	During the pretest, the use of mobile-based assessment and paper-based assessment modes had no effect on students' English vocabulary acquisition. However, after learning, students' vocabulary acquisition which was tested with a mobile application showed higher effectiveness compared to paper-based assessment.
Comparability of Computer-Based Testing and Paper-Based Testing: Testing Mode Effect, Testing Mode Order, Computer Attitudes and Testing Mode preference (Khoshshima & Toroujeni, 2017)	Most students show a preference for high CBT over PBT. The advantages of CBT are that item items are easier to read, options are easier to choose, answers can be changed easily, and scores can be seen directly. 70% of students like the CBT environment because it is more comfortable and reduces fatigue.
Students' Testing Preferences in Mathematics (Uukkivi et al., 2020)	The majority of respondents stated that web-based assessment is a good innovation. There was no statistically significant difference between students who preferred to take mathematics tests on the computer, or on paper, or did not mind the type of test given.
Students' Perception Towards National Examination 2017: Computer-Based Test or Paper-Based Test (Jamiludin et al., 2017)	High school students in Kendari City prefer the PBT system for the national final exam. Some of the disadvantages of the CBT system are: it takes more time to get used to students operating the computer, students look confused when using CBT, are very dependent on electricity and if there is a blackout, students have to work on their answers again, using the computer for a long time causes eye fatigue
Developing creativity and problem-solving skills of engineering students: a comparison of web- and pen-and-paper-based approaches (Valentine et al., 2017)	In the problem solving process, students produce solutions using paper-based templates or equivalent web interfaces.

Table 3. Result Analysis on second research Question (P2)

Title	Results
Effect of Goal Orientation on EFL Learners' Performances in CBT and PBT Across Gender (Jalali et al., 2014)	The difference in the average score of students taking CBT and PBT is 2.37. Higher scores are shown when students take exams with PBT. The tendency of students to use computers in the technological era does not have a positive effect on improving student performance in CBT.
Are the Keyboards Weightier than the Biro's? The Effect of Computer-Based Testing on Students' Achievement and Anxiety in Mathematics	There is no significant difference in the average score of mathematics learning outcomes tested using computer-based tests and paper-based tests.

Adeneye O. A. (Awofala, 2020)	CBT is more effective in measuring large samples with different cultural backgrounds
Investigating the Role of Formative Mobile Based Assessment in Vocabulary Learning of Pre-Intermediate EFL Learners in Comparison with Paper Based Assessment (Yarahmadzahi & Goodarzi, 2020)	There is no significant effect on the average score of students' English skills on CBT and PBT. However, the average score of students with CBT was slightly higher than that of PBT.
Comparability of Computer-Based Testing and Paper-Based Testing: Testing Mode Effect, Testing Mode Order, Computer Attitudes and Testing Mode preference (Khoshshima & Toroujeni, 2017)	The assessment system using CBT and PBT does not affect student performance when given alternately. There is no relationship between computer attitude and CBT performance.
Computer-Based Assessment: Its Effects on Students' Performance and Attitude (Sobremisana & Aragon, 2016)	The results show that better performance is supported by computer scoring environment compared to conventional environment

2. Discussion

a. Effectiveness *Computer-Based Test (CBT)* and *Paper-Based Test (PBT)*

Based on the results of a systematic review of eight studies examining the effectiveness of Paper-Based Tests (PBT) and Computer-Based Tests (CBT), it can be concluded that the effectiveness of assessment implementation is highly influenced by contextual factors, including implementation settings, infrastructure readiness, and learner characteristics. In general, CBT demonstrates advantages in terms of efficiency and user experience; however, it cannot be considered universally more effective across all contexts.

Several studies indicate that CBT has a stronger potential to enhance student motivation. Piaw (2012) reported that students tend to exhibit higher levels of motivation when engaging with CBT compared to PBT. Key motivational aspects such as involvement, curiosity, and self-efficacy are more effectively supported in CBT environments due to their interactive features and visually engaging interfaces. In addition, CBT can reduce student fatigue through concise item presentation, flexible testing time, and user-friendly navigation systems that facilitate easier interaction with test items.

Despite these advantages, the effectiveness of CBT is strongly dependent on technical readiness and students' digital literacy. Jamiludin et al. (2017) highlighted several challenges in CBT implementation, including the need for stable electricity, adequate hardware, and sufficient training to ensure students are familiar with computerized systems. In educational contexts where infrastructure is limited, PBT remains a more reliable and stable option. This finding is consistent with Awofala (2020), who emphasized that the effectiveness of CBT significantly declines when technical requirements are not adequately met.

Student preference also plays a crucial role in shaping perceptions of effectiveness. While many students prefer CBT due to its practicality and technological integration (Khoshshima & Toroujeni, 2017), others demonstrate resistance stemming from limited computer experience, discomfort with prolonged screen use, and concerns about potential technical disruptions during examinations (Hosseini et al., 2014).

From an administrative perspective, CBT offers considerable advantages, including automated scoring, rapid result generation, and efficient data storage. In contrast, PBT, although simple and widely accessible, requires more time for scoring and is more susceptible to human error.

From an analytical standpoint, these findings suggest that the effectiveness of assessment methods cannot be evaluated solely based on technological advancement. Instead, effectiveness should be understood as a multidimensional construct that encompasses pedagogical alignment, technological readiness, and user adaptability. CBT aligns well with contemporary educational paradigms that emphasize digital competence and efficiency; however, its implementation must be accompanied by adequate infrastructure and capacity building to ensure equitable access and valid measurement outcomes.

Furthermore, this review highlights the importance of adopting a context-sensitive approach in selecting assessment methods. Rather than positioning CBT as a replacement for PBT, both methods should be viewed as complementary tools within a flexible assessment framework. For instance, CBT may be more suitable for large-scale, standardized, and formative assessments, while PBT may remain advantageous in contexts requiring deep reading, extended responses, or where technological constraints persist.

The findings of this systematic review reveal a consistent pattern indicating that the effectiveness of assessment methods is not solely determined by the testing mode itself, but rather by the interaction between technological, pedagogical, and contextual factors. Across the reviewed studies, Computer-Based Testing (CBT) demonstrates notable advantages in terms of efficiency, user experience, and administrative practicality; however, these advantages are not universally applicable across all educational contexts.

A recurring trend observed in the literature is the positive influence of CBT on student motivation. Several studies report that students exhibit higher levels of engagement, curiosity, and self-efficacy when assessed through digital platforms. This can be explained by the interactive and visually stimulating nature of CBT environments, which align with contemporary motivational learning theories emphasizing the role of engagement and immediate feedback in enhancing learning experiences. Furthermore, CBT reduces physical fatigue associated with prolonged writing tasks and allows for more flexible navigation through test items, thereby improving the overall testing experience.

However, the effectiveness of CBT is highly contingent upon infrastructural readiness and students' digital literacy. The findings indicate that in contexts where technological resources are limited—such as unstable electricity, inadequate hardware, or insufficient training—CBT may become less effective or even counterproductive. This suggests that the benefits of CBT are conditional rather than inherent. From a systems perspective, the successful implementation of CBT requires alignment between technological infrastructure and user competence. Without such alignment, the intended advantages of digital assessment may not be realized.

b. Student Performance on PBT and CBT

The analysis of student performance across PBT and CBT reveals a notable convergence of findings, indicating that there is generally no statistically significant difference between the two assessment modes. This consistency across multiple studies suggests that the mode of test delivery does not directly influence students'

cognitive achievement. Instead, both PBT and CBT appear to provide equivalent measures of learning outcomes when implemented under comparable conditions.

However, a deeper examination of the literature reveals that variations in performance are influenced by several mediating factors, including students' familiarity with technology, test-taking behavior, and the nature of the assessment tasks. This indicates that student performance is not inherently determined by the testing medium, but rather by the interaction between the medium and learner characteristics.

In some cases, students demonstrate higher performance in PBT environments. This can be attributed to their greater familiarity with traditional testing formats, which allow for more flexible reading strategies, annotation, and cognitive processing. From the perspective of cognitive load theory, paper-based environments may reduce extraneous cognitive load for students who are less accustomed to digital interfaces, thereby facilitating better performance.

Conversely, other studies report improved performance in CBT settings, particularly in disciplines such as science and mathematics. This advantage can be explained by the affordances of digital platforms, including clearer graphical representations, interactive elements, and more efficient navigation. These features may enhance comprehension and problem-solving efficiency, especially for tasks that involve visual or numerical processing.

Although many studies report no significant differences in average scores, there is a tendency for CBT to produce slightly higher performance outcomes in certain contexts. This trend may be associated with increased efficiency in answering questions, ease of revising responses, and better time management. Students with higher levels of digital literacy are particularly advantaged in CBT environments, as they can utilize the interface more effectively.

Another critical factor influencing performance is test-taking behavior. While core strategies such as planning and evaluation are common across both modes, CBT facilitates additional strategies such as real-time monitoring, quick navigation, and answer revision. These affordances may contribute to improved performance by enabling more adaptive and efficient test-taking processes.

From an analytical standpoint, these findings reinforce the argument that performance differences between PBT and CBT are not rooted in the assessment medium itself, but in the broader interaction between technology, learner characteristics, and task demands. This highlights the importance of considering equity in assessment design, as disparities in digital access and literacy may create unintended advantages or disadvantages among students.

Furthermore, these results challenge the assumption that technological advancement inherently leads to improved learning outcomes. Instead, they suggest that the effectiveness of digital assessment depends on its alignment with learners' competencies and contextual conditions. Thus, the equivalence observed between PBT and CBT should be interpreted as evidence of measurement consistency rather than technological superiority.

In addressing the second research question (P2), it can be concluded that student performance is generally comparable across PBT and CBT. However, variations may occur depending on factors such as digital familiarity, subject characteristics, and test-taking strategies. Therefore, the choice of assessment mode should be guided by pedagogical considerations and contextual readiness, rather than assumptions about inherent performance advantages.

D. CONCLUSION AND SUGGESTIONS

The review indicates that CBT and PBT generally produce comparable performance outcomes, but equivalence is conditional on learner characteristics, task demands, interface design, and implementation context. CBT frequently provides advantages in motivation, processing speed, scoring, and administration; however, these benefits should not be interpreted as evidence that CBT is inherently more valid or produces higher achievement. PBT may remain advantageous for learners who rely on familiar paper-based strategies or study in settings with limited technological access.

Educational institutions should select assessment modes through evidence-based pilot testing and provide adequate digital preparation, accessible interfaces, reliable infrastructure, and contingency procedures. Future comparative studies should prioritize score equivalence, measurement validity, and fairness across diverse learner groups rather than focusing solely on preference or administrative efficiency.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support provided by the Faculty of Teacher Training and Education, Universitas Sulawesi Barat, in facilitating this research. The authors also extend their appreciation to colleagues and academic peers for their valuable discussions and feedback throughout the research process.

REFERENCES

- Awofala, A. O. A. (2020). Indonesian Journal of Informatics Education Are the Keyboards Weightier than the Biro's? The Effect of Computer-Based Testing on Students' Achievement and Anxiety in Mathematics. *Indonesian Journal of Informatics Education*, 4(1), 1–8.
- Bhoola-Patel, A., & Laher, S. (2011). The Influence of Mode of Test Administration on Test Performance. *Journal of Psychology in Africa*, 21(1), 139–144. <https://doi.org/10.1080/14330237.2011.10820440>
- Brown, G. T. L. (2019). Technologies and infrastructure: costs and obstacles in developing large-scale computer-based testing. *Education Inquiry*, 10(1), 4–20. <https://doi.org/10.1080/20004508.2018.1529528>
- Budiastuti, P., Soenarto, S., Muchlas, M., & Ramndani, H. W. (2021). Analysis of Learning Objectives and Basic Competencies in the Lesson Plan for Basic Electricity and Electronics at Vocational High Schools. *Jurnal Edukasi Elektro*, 5(1), 39–48. <https://doi.org/10.21831/jee.v5i1.37776>
- Carpenter, R., & Alloway, T. (2019). Computer Versus Paper-Based Testing: Are They Equivalent When it Comes to Working Memory? *Journal of Psychoeducational Assessment*, 37(3), 382–394. <https://doi.org/10.1177/0734282918761496>
- Ferdiansyah, S. (2016). Moving from Paper-Based Testing (PBT) to Computer-Based Testing (CBT) for Classroom Use: Exploring the Opportunities and Challenges. *Prosiding ICTTE FKIP UNS 2015*, 1, 514–519.
- García-Pérez, L., García-Garnica, M., & Olmedo-Moreno, E. M. (2021). Skills for a working future: How to bring about professional success from the educational setting. *Education Sciences*, 11(1), 1–25. <https://doi.org/10.3390/educsci11010027>
- Hakim, B. M. (2018). Comparative Study on Validity of Paper-Based Test and Computer-Based Test in the Context of Educational and Psychological Assessment among Arab Students. 8(2), 85–91. <https://doi.org/10.5539/ijel.v8n2p85>
- Hassler Hallstedt, M., & Ghaderi, A. (2018). Tablets instead of paper-based tests for young children? Comparability between paper and tablet versions of the mathematical Heidelberg Rechen Test 1-4. *Educational Assessment*, 23(3), 195–210. <https://doi.org/10.1080/10627197.2018.1488587>
- Hosseini, M., Jafre, M., Abidin, Z., & Baghdarnia, M. (2014). Comparability of Test Results of Computer Based Tests (CBT) and Paper and Pencil Tests (PPT) among English Language

- Learners in Iran. *Procedia - Social and Behavioral Sciences*, 98, 659–667. <https://doi.org/10.1016/j.sbspro.2014.03.465>
- Jalali, S., Zeinali, M., & Nobakht, A. (2014). Effect of Goal Orientation on EFL Learners' Performances in CBT and PBT Across Gender. *Procedia - Social and Behavioral Sciences*, 98, 727–734. <https://doi.org/10.1016/j.sbspro.2014.03.474>
- Jamiludin, Darnawati, & Uke, W. A. S. U. (2017). Students' Perception Towards National Examination 2017: Computer-Based Test or Paper-Based Test. *Mediterranean Journal of Social Sciences*, 8(4), 139–144.
- Karay, Y., Schaubert, S. K., Stodch, C., & Schüttelz-Brauns, K. (2015). Computer Versus Paper — Does It Make Any Difference in Test Performance? *Teaching and Learning in Medicine*, 27(1), 57–62. <https://doi.org/10.1080/10401334.2014.979175>
- Khoshshima, H., & Toroujeni, S. M. H. (2017). Comparability of Computer-Based Testing and Paper-Based Testing: Testing Comparability of Computer-Based Testing and Paper-Based Testing: Testing Mode Effect, Testing Mode Order, Computer Attitudes and Testing Mode preference. *International Journal of Computer*, 24(1), 80–99.
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51(1), 7–15. <https://doi.org/10.1016/j.infsof.2008.09.009>
- Munna, A. S., & Kalam, M. A. (2021). Teaching and learning process to enhance teaching effectiveness: literature review. *International Journal of Humanities and Innovation (IJHI)*, 4(1), 1–4. <https://doi.org/10.33750/ijhi.v4i1.102>
- Piaw, C. Y. (2012). Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong? *Procedia-Social and Behavioral Sciences*, 64, 655–664. <https://doi.org/10.1016/j.sbspro.2012.11.077>
- Sobremisana, V. S., & Aragon, R. C. (2016). Computer-Based Assessment: Its Effects On Students' Performance And Computer-Based Assessment: Its Effects On Students' Performance And Attitude. *The International Journal Research Publication's*, 5(9), 10–18.
- Torres-Carrión, Vicente, P., Aciar, Silvana, González-González, Soledad, C., Rodríguez-Morales, & Germania. (2018). Methodology for Systematic Literature Review applied to Engineering and Education. *IEEE Global Engineering Education Conference (EDUCON)*, 1364–1373.
- Uukkivi, A., Labanova, O., Retshnoi, V., Safiulina, E., & Latõnina, M. (2020). Students' Testing Preferences in Mathematics. *2nd International Conference on Research in Teaching and Education*, 63–73.
- Valentine, A., Belski, I., Hamilton, M., Valentine, A., Belski, I., Developing, M. H., & Valentine, A. (2017). Developing creativity and problem-solving skills of engineering students: a comparison of web- and pen-and-paper-based approaches. *European Journal of Engineering Education*, 0(0), 1–21. <https://doi.org/10.1080/03043797.2017.1291584>
- Yanti, M., Riandi, & Suhandi, A. (2019). Analyzing Tpack Ability of Science Teacher Based on Experience for Teaching Global Warming in Secondary Level. *Unnes Science Education Journal*, 8(2), 130–138.
- Yarahmadzahi, N., & Goodarzi, M. (2020). Investigating The Role of Formative Mobile Based Assessment in Vocabulary Learning of Pre-Intermediate EFL Learners in Comparison with Paper Based Assessment. *Turkish Online Journal of Distance Education*, 21(1), 181–196.
- Zoghi, M. (2013). *Computer-based vs . Paper-based Writing Tests: A Comparative Study of Test Taking Strategies*. 1(3), 43–56. <https://doi.org/10.5296/ijele.v1i3.3559>