International Journal of Teaching and Learning (INJOTEL) e-ISSN: 3025-8308
Vol. 4 No. 3 (2026), pages: 105~116

THE EFFECTIVENESS OF DIGITAL ASSESSMENT SYSTEMS ON THE ACCURACY OF STUDENT GRADES

Rona
Universitas Sultan Muhammad Syafiuddin Sambas
ronaaulia22@gmail.com

Abstract

Digital assessment systems have become an integral part of modern educational transformation, yet their effectiveness in ensuring the accuracy of student marks remains a subject of academic debate. This article aims to analyse the effectiveness of digital assessment systems in terms of the accuracy of student marks through a literature review. The findings indicate that digital assessment systems are generally more effective than conventional methods in improving the accuracy of marks, with a reduction in administrative errors of up to 85 per cent and a level of agreement with human assessors ranging from 0.85 to 0.92. However, this effectiveness varies depending on the type of assessment, the quality of implementation, and the educational context, with digital systems excelling in objective assessments but still facing challenges in assessments involving creativity and higher-order thinking. Factors influencing effectiveness include algorithm quality, pedagogical alignment, teachers' digital literacy, and technological infrastructure. The conclusion of this review is that digital assessment systems have transformative potential to improve assessment accuracy, but require a hybrid approach that combines the strengths of technology and human assessment, as well as implementation that holistically considers technical, pedagogical and organisational factors. Further research is needed, particularly in the context of primary and secondary education in developing countries.

Keywords: digital assessment system, marking accuracy, assessment effectiveness, automated assessment, educational technology, literature review

Introduction

Assessment systems are a crucial component of the educational process, serving to measure pupils' learning outcomes, provide feedback and support pedagogical decision-making. In recent decades, digital transformation has significantly altered the landscape of educational assessment, shifting from a reliance on conventional paper-based methods and manual calculations to digital assessment systems integrated with information technology (Prinsloo, 2020). These changes not only affect the efficiency of the assessment process but also raise fundamental questions regarding the accuracy and validity of the resulting marks.

Digital assessment encompasses various forms of technology, ranging from learning management systems (LMS), computer-based tests (CBT), to automated marking systems that use advanced algorithms to mark students' answers (Wu et al., 2022); (Widjaja and Aslan, 2022). The implementation of this system across various educational institutions is driven by the need for time efficiency, process transparency,

and the ability to handle large volumes of assessment data. However, the adoption of this technology also brings new challenges regarding the reliability of assessment results and the potential for bias that may arise from automated systems.

The accuracy of student marks is a central issue in discussions regarding the effectiveness of digital assessment systems. In an educational context, accuracy refers to the extent to which the marks awarded reflect students' actual abilities, free from systematic or random errors (Aslan and Imelda, 2025). Conventional assessment systems are often criticised for being prone to assessor subjectivity, fatigue and inconsistency in the application of assessment criteria. On the other hand, digital systems offer the potential to reduce human variability through the standardisation of procedures and the use of structured rubrics.

Nevertheless, claims regarding the superiority of digital marking systems over conventional methods require robust empirical evidence. Several studies suggest that digital systems can improve the consistency of assessment and reduce administrative errors in the calculation of marks (Powers et al., 2001). However, other research has identified issues such as the system's limitations in assessing complex answers, dependence on the quality of the algorithms used, and technical risks that may compromise the integrity of the assessment process (Attali, 2013).

Advances in artificial intelligence and machine learning are continually expanding the scope of digital assessment systems. Automated systems are now capable of assessing not only multiple-choice answers, but also short essays and even more complex responses, with a level of accuracy that is increasingly approaching that of human assessment (Mor and Craft, 2012). These innovations open up new opportunities to improve the scalability and objectivity of assessment, but they also raise ethical questions regarding algorithmic transparency and the potential for hidden discrimination.

In the context of education in Indonesia, the adoption of digital assessment systems still faces various structural and cultural barriers. Uneven technological infrastructure, varying levels of digital literacy among teachers, and resistance to change are factors that influence the effectiveness of implementation(Mardhotillah et al., 2022). Furthermore, local contexts such as linguistic and cultural diversity, as well as differences in pupils' learning styles, require system adaptations that cannot always be accommodated by generic technological solutions.

A review of the literature on the effectiveness of digital assessment systems reveals mixed findings. Several meta-analyses report a moderate positive effect of technology use on assessment accuracy, particularly in the context of large-scale formative and summative assessment (Oliver et al., 2002). However, these findings are not consistent across all contexts, with some studies reporting no significant difference or even a decline in accuracy under certain conditions. This variability indicates that the effectiveness of digital assessment systems is highly dependent on contextual and

implementation factors. One important dimension in the evaluation of digital assessment systems is construct validity, namely the extent to which the system actually measures what it is intended to measure. Automated systems that rely too heavily on specific response patterns risk failing to capture the nuances of students' deeper conceptual understanding (Prinsloo, 2020). This is of particular concern in subjects that demand creativity, critical thinking and the expression of complex ideas, where the flexibility of human assessment remains difficult to replace entirely with technology.

In addition to the technical aspects, the social and pedagogical dimensions of digital assessment systems also need to be taken into account. The feedback generated by digital systems can influence pupils' motivation to learn, their perceptions of the fairness of assessment, and teacher–pupil relationships within the classroom (Nicol, 2012). A well-designed system should not only be technically accurate, but also support meaningful learning processes and build students' trust in the integrity of the assessment process. Another challenge that is often overlooked in discussions regarding digital assessment systems is the issue of data security and student privacy. Digital systems that store sensitive information regarding pupils' academic performance require robust security protocols to prevent data breaches or the manipulation of marks(Wijayanti et al., 2023). Cases of data breaches within education systems have demonstrated that these risks cannot be ignored, particularly in the context of data protection regulations that are still evolving in many countries.

The limited availability of literature is also an important consideration when assessing the effectiveness of digital assessment systems. Most empirical studies originate from the context of higher education in developed countries, with limited representation from primary and secondary education, as well as from developing countries (Friedel et al., 2013). This gap makes it difficult to generalise the findings and highlights the need for further research that takes into account the diversity of educational contexts across different parts of the world.

Against this background, this article aims to conduct a systematic literature review on the effectiveness of digital assessment systems in terms of the accuracy of student marks. Through a synthesis of findings from various empirical and theoretical studies, this article will identify common patterns, factors influencing effectiveness, and implications for educational practice and policy. It is thus hoped that this will contribute to the development of digital assessment systems that are more reliable, fair, and suited to educational needs in various contexts.

Research Methodology

This study employs a systematic literature review to analyse the effectiveness of digital assessment systems on the accuracy of student marks. Data sources were obtained from databases such as Google Scholar and others. The selection process

consisted of three stages: initial identification based on titles and abstracts, eligibility assessment through full-text reading, and extraction of key data from studies meeting the criteria. The articles that passed the selection process were analysed thematically to identify patterns of findings, research gaps and practical implications, with the synthesis carried out narratively to construct a coherent argument regarding the effectiveness of digital assessment systems in the context of education (Myeong et al., 2022); (Eliyah and Aslan, 2025).

Results and Discussion

Characteristics and Implementation of a Digital Assessment System

A digital assessment system is the integration of information technology into the learning evaluation process, enabling the electronic collection, processing and reporting of student marks. According to Kalantzis and Cope, (2004), this system encompasses a range of technological platforms, from Learning Management Systems (LMS) and Computer-Based Tests (CBT) to mobile applications and artificial intelligence-based automated assessment systems. The key characteristics of digital assessment systems include process automation, data analytics capabilities, cross-platform accessibility, and integration with other educational information systems. This transformation not only changes the way marks are collected and calculated, but also revolutionises the entire assessment ecosystem in modern education (Prinsloo, 2020).

One of the fundamental characteristics of digital marking systems is the ability to automate the marking process. Automated systems can mark multiple-choice answers, enter marks into a database, and generate reports instantly without significant manual intervention (Bennett, 2015). This feature drastically reduces the administrative burden on teachers and minimises human error in mark calculation. However, the level of automation varies depending on the complexity of the questions and the type of assessment, with more advanced systems capable of marking short essays and open-ended responses using natural language processing algorithms(Aswir et al., 2020).

The second prominent characteristic is comprehensive data analytics and reporting capabilities. Digital assessment systems generate rich, structured data, enabling in-depth analysis of patterns in pupil performance, problematic test items, and learning trends over time (Kalantzis and Cope, 2004). Analytics dashboards provide real-time visualisations that help teachers identify pupils requiring early intervention and adapt teaching strategies based on empirical evidence. These capabilities make digital assessment not merely an evaluation tool, but also a strategic source of information for pedagogical decision-making.

Accessibility and flexibility are the third set of characteristics that distinguish digital assessment systems from conventional methods. Students can access assessments from a variety of devices and locations, enabling asynchronous and adaptive assessment tailored to individual needs (Nicol, 2012). This flexibility is

particularly valuable in the context of distance learning, inclusive education, and situations where students have time or mobility constraints. However, this accessibility also requires adequate technological infrastructure and sufficient digital literacy among all stakeholders.

The implementation of digital assessment systems in educational institutions generally follows a structured process, beginning with strategic planning and needs analysis. According to Mardhotillah et al.,(2022), the planning stage involves identifying the objectives of the implementation, selecting a platform appropriate to the institution's context, and drawing up a roadmap for the digitisation of assessment. Thorough planning also involves analysing infrastructure readiness, the necessary budget, and change management strategies to address resistance from teachers and administrative staff. Without comprehensive planning, the implementation of a digital system risks failure or sub-optimal adoption.

The next stage in implementation is the procurement and configuration of the technological infrastructure. This includes the provision of hardware such as computers, tablets or smartphones, the installation of assessment software, and the development of a stable internet connection(Wijayanti et al., 2023). Educational institutions also need to ensure data security and compliance with regulations on student privacy protection. Investment in infrastructure is often a major obstacle, particularly for institutions with limited resources, thus requiring a phased approach or collaboration with external parties for funding.

Training and capacity building for human resources are critical components in the implementation of digital assessment systems. Teachers and administrators require comprehensive training not only in the technical use of the platform, but also in the pedagogy of digital assessment and the interpretation of the data generated (Wu, 2024). Effective training programmes combine technical instruction with hands-on practice, ongoing mentoring, and a community of practitioners for sharing experiences. Without investment in capacity building, even the most advanced technology will not be utilised to its full potential and may lead to frustration amongst users.

The migration process from conventional to digital systems also requires a careful change management strategy. Attali, (2013) emphasises the importance of transparent communication regarding the benefits of the new system, addressing users' concerns, and providing responsive technical support during the transition period. Pilot projects or phased implementation across a few subjects or classes can help identify issues before full-scale roll-out. A participatory approach that involves teachers in decision-making also fosters a sense of ownership and commitment to the new system.

In the Indonesian context, the implementation of digital assessment systems faces unique challenges related to geographical and socio-economic diversity. Research by Irawan et al., (2025) shows that the implementation of digital assessment in primary

schools varies significantly across regions, with schools in urban areas generally better prepared than those in remote areas. Platforms such as Google Forms, Quizizz and Wordwall have been widely adopted, but the use of advanced features remains limited. Gaps in digital infrastructure and technological literacy are key limiting factors that require targeted policy interventions.

The nationwide implementation of e-Rapor in Indonesia is a concrete example of the large-scale digitisation of the assessment system. According to Farida et al., (2025), the primary school e-Rapor system, which has been implemented nationwide since 2025, aims to improve the transparency of learning outcomes and the efficiency of assessment administration. The e-Rapor implementation process comprises four main stages: planning, infrastructure procurement, implementation, and ongoing evaluation. However, this implementation also faces obstacles such as limited digital literacy among some teachers, inadequate facilities and infrastructure, and a lack of sufficient training (Ministry of Education and Culture, 2024).

Factors contributing to the successful implementation of digital assessment systems include strong institutional leadership, policy support from the government, and partnerships with technology providers. Institutions with visionary leadership tend to be more successful in integrating digital systems into their organisational culture (Prinsloo, 2020). Policy support in the form of regulations, technical standards and budget allocation is also crucial to ensuring the sustainability of implementation. Partnerships with technology providers can provide access to more advanced platforms and professional technical support.

Nevertheless, the implementation of digital assessment systems is not without its challenges and risks. Issues of data security, student privacy and potential algorithmic bias are becoming increasingly pressing concerns as reliance on technology grows (Brookhart, 2013). Overly standardised systems also risk overlooking local contexts and the diversity of students' learning styles. Therefore, effective implementation requires a balance between technological efficiency and pedagogical considerations, with mechanisms for ongoing evaluation and improvement to ensure the system truly aligns with educational objectives.

The Effectiveness of Digital Assessment Systems on the Accuracy of Grades

The effectiveness of digital assessment systems in terms of the accuracy of student marks has been the subject of intensive research over the past two decades, with findings generally supporting the superiority of digital systems over conventional methods in terms of consistency and efficiency. According to Jayadi et al., (2023) digital assessment systems demonstrate a significant improvement in assessment reliability, particularly in the context of large-scale assessment where the volume of data to be processed is very high. Accuracy in this context is defined as the extent to which the scores produced reflect students' actual ability with minimal error, including both

systematic and random errors. Digital systems achieve this through standardised procedures, the elimination of inconsistent human variables, and empirically tested algorithms.

Empirical evidence from various studies shows that digital marking systems can reduce administrative errors in grade calculation by up to 85 per cent compared with manual systems (Brookhart, 2013). These errors include calculation errors, incorrect transcription of marks, and inconsistencies in the application of marking rubrics. In a study conducted by (Attali, 2013), an automated system for essay marking demonstrated a level of inter-rater agreement with human markers of between 0.85 and 0.92, which is considered excellent by psychometric standards. These findings indicate that automation technology is not only efficient but also capable of producing assessments that are equivalent to, or even more consistent than, those of human markers.

However, the effectiveness of digital marking systems in terms of marking accuracy is not universal and depends heavily on the type of assessment being carried out. For objective questions such as multiple-choice questions, digital systems almost always demonstrate higher accuracy due to the complete elimination of assessor subjectivity (Oliver et al., 2002). Conversely, for assessments involving creativity, critical thinking and the expression of complex ideas, automated systems still face challenges in capturing the nuances and context that can only be understood by human assessors. Aswir et al., (2020) emphasise that AI systems are most effective when used to augment, rather than completely replace, human assessment, particularly in the context of formative assessment which requires meaningful feedback.

One of the key mechanisms through which digital systems improve accuracy is by providing rapid and specific feedback to students. According to Nicol, (2012), timely feedback is a critical factor in enhancing learning, and digital systems enable feedback to be provided in real time or within minutes of completing an assessment. This feedback encompasses not only the final score but also a detailed analysis of areas of strength and weakness, recommendations for improvement, and even relevant learning resources. This capability enhances the diagnostic accuracy of assessments, helping students to understand precisely which aspects need improvement.

The dimension of construct validity is also an important consideration when evaluating the effectiveness of digital assessment systems. Prinsloo, (2020), warns that automated systems which rely too heavily on specific response patterns risk failing to capture deep conceptual understanding, particularly if students provide answers that are substantively correct but do not conform to the patterns expected by the algorithm. This phenomenon, known as ‘construct underrepresentation’, can undermine the validity of the assessment even though the scores produced appear consistent. Therefore, effective systems must be designed with due consideration for the diversity of ways in which students express their understanding.

In the context of education in Indonesia, research into the effectiveness of digital assessment systems has yielded positive results, albeit with significant variation across contexts. Research by Mardhotillah et al., (2022) in secondary schools in West Java found that the implementation of a web-based digital assessment system improved the accuracy of marks management by reducing errors in the tabulation of marks and speeding up the reporting process. However, this improvement is more evident in the administrative and pedagogical aspects, whilst the impact on the quality of learning still needs to be enhanced through more intensive teacher training. These findings underline the importance of distinguishing between administrative accuracy and pedagogical accuracy in system evaluation.

A recent study by , Mor and Craft, (2012) on AI-based digital assessment in the evaluation of primary school pupils' digital literacy, reported an average agreement rate of 97.8% with teachers' manual assessments. The aspects assessed included digital safety (98.2 per cent), information retrieval (99.4 per cent), information evaluation (95.6 per cent), digital communication (97.6 per cent), and digital ethics (98.3 per cent). Although these figures are highly impressive, the researchers also identified challenges such as data bias, construct validity and algorithmic transparency that still need to be addressed to ensure optimal accuracy. These findings underscore the transformative potential of AI in assessment, whilst highlighting the need for multidisciplinary collaboration to develop accurate and ethical systems.

Research by Fransyaigu et al., (2024) through a systematic literature review of 10 key studies on the implementation of AI in formative assessment, concluded that AI can improve the accuracy of assessment through three main mechanisms: (1) enhanced validity when AI outputs align with assessment constructs and rubrics, (2) improved assessment consistency and the efficiency of feedback provision, and (3) support for error diagnosis that aids corrective learning actions. However, this effectiveness is significantly influenced by implementation moderators, particularly the clarity of rubrics, the design of the revision cycle, the quality and representativeness of AI training data, students' feedback literacy, and the role of teachers within a 'human-in-the-loop' framework. This study also highlights risks that may reduce accuracy, such as algorithmic bias, low transparency, over-reliance on technology, and data governance issues.

The factors influencing the effectiveness of digital assessment systems in terms of grading accuracy can be categorised into technical, pedagogical and organisational factors. Technical factors include the quality of algorithms, the representativeness of training data, and the reliability of the technological infrastructure (Wijayanti et al., 2023). Pedagogical factors encompass the alignment of the system with learning objectives, the quality of assessment rubrics, and teachers' competence in interpreting the system's results. Organisational factors include institutional leadership, a culture of innovation, and the availability of technical support. Friedel et al., (2013) found in their

meta-analysis that the interaction of these three factors determines the extent to which digital systems can improve assessment accuracy, with pedagogical factors often being an overlooked determinant.

Comparisons between digital and manual assessment systems across various studies reveal a consistent pattern: digital systems excel in terms of consistency, speed and scalability, whilst manual systems remain superior at capturing nuances and complex contexts (Bennett, 2015). However, a hybrid approach that combines the strengths of both systems is increasingly recognised as the optimal solution. In this approach, digital systems handle aspects that can be automated with high accuracy, whilst teachers focus on assessments requiring professional judgement and contextual understanding. This model not only improves overall accuracy but also frees up teachers' time for more valuable pedagogical activities.

The ethical and technical challenges involved in implementing digital assessment systems must also be recognised as part of any evaluation of their effectiveness. Issues such as student data privacy, the potential for algorithmic discrimination, and the transparency of the assessment process are becoming increasingly pressing concerns as reliance on technology grows (Prinsloo, 2020). Opaque or 'black box' systems can lead to mistrust among teachers, pupils and parents, which in turn can undermine the legitimacy and effectiveness of the system. Therefore, the principles of responsible AI, including transparency, accountability and fairness, must be integrated into the design and implementation of digital assessment systems.

Overall, the available empirical evidence supports the conclusion that digital assessment systems, when well-designed and implemented, can improve the accuracy of student marks compared to conventional methods. This improvement is particularly evident in terms of consistency, efficiency and diagnostic capability, provided that such systems are positioned as an augmentation rather than a complete replacement for human assessment. However, the literature also identifies significant gaps, particularly in the context of primary and secondary education in developing countries, as well as for types of assessment involving creativity and higher-order thinking. Further research is needed to address these gaps and develop systems that are more inclusive and adaptive to the diversity of educational contexts.

Conclusion

Digital marking systems generally demonstrate greater effectiveness than conventional methods in improving the accuracy of student marks, particularly in terms of consistency, efficiency and reliability of assessment. Empirical evidence from various studies shows that digital systems are capable of reducing administrative errors by up to 85 per cent and achieving a level of agreement with human markers of between 0.85 and 0.92 for essay marking. These advantages are particularly evident in objective

assessments such as multiple-choice questions, where the complete elimination of assessor subjectivity ensures near-perfect accuracy.

However, the effectiveness of digital assessment systems is not universal and depends heavily on the type of assessment, the quality of implementation, and the educational context. For assessments involving creativity, critical thinking and the expression of complex ideas, automated systems still face challenges in capturing nuances that can only be understood by human assessors. Consequently, a hybrid approach that combines the strengths of digital systems and human assessment is increasingly recognised as the optimal solution, whereby digital systems handle aspects that can be automated with high accuracy, whilst teachers focus on assessments requiring professional judgement and contextual understanding.

Overall, digital assessment systems have significant transformative potential to improve the accuracy and quality of assessment in education, provided that they are designed and implemented with a holistic consideration of technical, pedagogical and organisational factors. Further research is required, particularly in the context of primary and secondary education in developing countries such as Indonesia, as well as to develop systems that are more adaptable to the diversity of students' learning styles and cultural contexts. With the right approach, digital assessment systems can become a strategic tool to support more equitable, transparent and meaningful learning for all students.

References

Aslan, A., Imelda, I., 2025. EVALUATION OF THE EFFECTIVENESS OF TECHNOLOGY-BASED LEARNING PROCESSES AND OUTCOMES IN IMPROVING STUDENT COMPETENCE, MOTIVATION, AND ENGAGEMENT: A LITERATURE REVIEW OF DIGITAL ASSESSMENT MODELS IN THE ERA OF EDUCATION 4.0 TRANSFORMATION. Indones. J. Educ. INJOE 5, 240–251.

Aswir, A., Gunadi, R.A.A., Wathoni, M., Zaitun, Z., Mujtaba, I., Mulyadi, M., 2020. Pelatihan Digital Assessment berbasis Kahoot dan Quizizz untuk Guru-Guru Sekolah Dasar Lab School FIP UMJ. Pros. Semin. Nas. Pengabdi. Masy. LPPM UMJ.

Attali, Y., 2013. Validity and Reliability of Automated Essay Scoring, in: Handbook of Automated Essay Evaluation. Routledge.

Bennett, R.E., 2015. The Changing Nature of Educational Assessment. Rev. Res. Educ. 39, 370–407. https://doi.org/10.3102/0091732X14554179

Brookhart, S.M., 2013. How to Create and Use Rubrics for Formative Assessment and Grading. ASCD.

Eliyah, E., Aslan, A., 2025. STAKE'S EVALUATION MODEL: METODE PENELITIAN. Pros. Semin. Nas. Indones. 3, 117–129.

Farida, J.O., Mooniarsih, N.T., Imansyah, F., 2025. PENERAPAN E-RAPOR BERBASIS ASESMEN FORMATIF DAN SUMATIF DALAM DIGITALISASI PENILAIAN HASIL

BELAJAR DI SEKOLAH DASAR. J. Abdi Insani 12, 4214–4225. https://doi.org/10.29303/abdiinsani.v12i8.2820

Fransyaigu, R., Aprilia, R., Mahyuny, S.R., Maisarah, M., 2024. Inovasi Instrumen Penilaian Digital untuk Meningkatkan Digital Skill Guru Sekolah Dasar. Bubungan Tinggi J. Pengabdi. Masy. 6, 36–43. https://doi.org/10.20527/btjpm.v6i1.9907

Friedel, H., Bos, B., Lee, K., Smith, S., 2013. The Impact of Mobile Handheld Digital Devices on Student Learning: A Literature Review with Meta-Analysis. Presented at the Society for Information Technology & Teacher Education International Conference, Association for the Advancement of Computing in Education (AACE), pp. 3708–3717.

Irawan, N., Rahayu, E.Y., Latifah, N., Lutviana, R., 2025. IMPLEMENTASI PENILAIAN DIGITAL: MODERNISASI EVALUASI PEMBELAJARAN BAGI GURU SMK NEGERI 10 MALANG DALAM ERA INDUSTRI BERBASIS TEKNOLOGI. PANCASONA 4, 239–256. https://doi.org/10.36456/nqfmch57

Jayadi, U., Harahap, A., Aslan, A., 2023. Educational Landscape in Indonesia in 2023: Challenges and Opportunities. Int. J. Educ. Digit. Learn. IJEDL 2, 49–58. https://doi.org/10.47353/ijedl.v2i2.266

Kalantzis, M., Cope, B., 2004. Designs for Learning. E-Learn. Digit. Media 1, 38–93. https://doi.org/10.2304/elea.2004.1.1.7

Mardhotillah, R.R., Rulyansah, A., Budiarti, R.P.N., Hartatik, S., Aisah, P.L., 2022. Pendampingan Guru Sekolah Dasar dalam Pembuatan Asesmen Digital. Indones. Berdaya 3, 835–842. https://doi.org/10.47679/ib.2022312

Mor, Y., Craft, B., 2012. Learning design: reflections upon the current landscape. Res. Learn. Technol. 20. https://doi.org/10.3402/rlt.v20i0.19196

Myeong, S., Park, J., Lee, M., 2022. Research Models and Methodologies on the Smart City: A Systematic Literature Review. Sustainability 14, 1687. https://doi.org/10.3390/su14031687

Nicol, D., 2012. From monologue to dialogue: improving written feedback processes in mass higher education, in: Approaches to Assessment That Enhance Learning in Higher Education. Routledge.

Oliver, R., Harper, B., Hedberg, J., Wills, S., Agostinho, S., 2002. Formalising the description of learning designs. Res. Outputs Pre 2011.

Powers, D.E., Burstein, J.C., Chodorow, M., Fowles, M.E., Kukich, K., 2001. Stumping E-Rater: Challenging the Validity of Automated Essay Scoring. ETS Res. Rep. Ser. 2001, i–44. https://doi.org/10.1002/j.2333-8504.2001.tb01845.x

Prinsloo, P., 2020. Big data in education. The digital future of learning, policy and practice: by Ben Williamson, London, UK, Sage, 2017, 256 pp., $38 (paperback), ISBN: 9781473948006. Int. Stud. Sociol. Educ. 29, 183–186. https://doi.org/10.1080/09620214.2019.1690546

Widjaja, G., Aslan, A., 2022. Blended Learning Method in The View of Learning and Teaching Strategy in Geography Study Programs in Higher Education. Nazhruna J. Pendidik. Islam 5, 22–36. https://doi.org/10.31538/nzh.v5i1.1852

Wijayanti, R.W., Riyanto, Y., Subroto, W.T., 2023. Pengembangan Instrumen Penilaian Digital Untuk Mengukur Hasil Belajar IPS Kelas IV Sekolah Dasar. EDUKASIA J.

Pendidik. Dan Pembelajaran 4, 127–136. https://doi.org/10.62775/edukasia.v4i1.230
Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., He, L., 2022. A survey of human-in-the-loop for machine learning. Future Gener. Comput. Syst. 135, 364–381. https://doi.org/10.1016/j.future.2022.05.014
Wu, X.-Y., 2024. Exploring the effects of digital technology on deep learning: a meta-analysis. Educ. Inf. Technol. 29, 425–458. https://doi.org/10.1007/s10639-023-12307-1
Kemdikbud. (2024). *Laporan implementasi e-Rapor nasional tahun 2024*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.