Evaluating the practicality of a HOTS assessment instrument in Pancasila education through the system usability scale

Dewiyana Agustina [1], Ai Sutini [2,a], Yunus Abidin [3,b], Dinie Anggraeni Dewi [4,c], Nissa Aulia Belistiana Utami [5,d]

[1,2,3,4] Universitas Pendidikan Indonesia, Indonesia

[5] Hiroshima University, Japan

[a] https://orcid.org/0009-0009-0534-5765

[b] https://orcid.org/0000-0003-3626-7404

[c] https://orcid.org/0000-0002-2716-2021

[d] https://orcid.org/0009-0005-3941-9241

✉ dewiyanaagustina@upi.edu

Abstract: This study aims to assess the practicality of a Higher Order Thinking Skills (HOTS) assessment instrument in Pancasila Education using the System Usability Scale (SUS) approach. A descriptive quantitative research method was employed, involving 30 fifth-grade primary school teachers of Pancasila Education. The practicality of the developed instrument was evaluated based on the SUS score and interpreted using standard usability criteria. The findings reveal that the instrument obtained an average SUS score of 76.5, which falls within the "practical" and "good" categories. Further analysis indicates that teachers found the instrument easy to use, the instructions clear, and felt confident applying it in classroom settings. These results underscore the importance of usability testing in the development of educational instruments, particularly in supporting 21st-century skills. The study recommends incorporating practicality testing as a key component of instrument validation and highlights the need for further research exploring usability in digital and longitudinal learning contexts.

Keywords: instrument practicality; HOTS; Pancasila education; System Usability Scale; usability testing.

Introduction

Global development in the 21st century introduces significant transformations across multiple dimensions of life, particularly within the realm of education. The dynamics of the Fourth Industrial Revolution, globalisation, and the era of digitalisation necessitate adaptive and creative human resources capable of critical thinking, collaboration, and effective communication (Trilling & Fadel, 2009). Education plays a strategic role in equipping learners with both mastery of knowledge content and the higher-order thinking skills necessary for future social, economic, and cultural engagement (Fullan & Scott, 2014). In the 21st century, education must adapt to keep pace with the rapidly evolving world (Mahmud & Wong, 2022; Shields & Chugh, 2018). This means accepting new technologies, teaching important skills, and constantly changing how schools work to make sure everyone has equal access to education, that it lasts, and that it is strong (Adler-Beléndez et al., 2020; Astuti et al., 2019; Mahmud & Wong, 2022; Shields & Chugh, 2018).

21st-century skills encompass critical thinking, complex problem-solving, creativity, leadership, collaboration, and digital literacy (Kocak et al., 2021; van Laar et al., 2020). These competencies cannot be developed through traditional rote learning; rather, they require active, collaborative, and problem-based learning approaches that cultivate students' higher-order thinking skills (Bhardwaj et al., 2025). In this context, Indonesia has implemented various educational policies, including the 2013 Curriculum and the Merdeka Curriculum, which emphasise the necessity of strengthening higher-order thinking competencies. Pancasila education, a core subject within the national education framework, aims not only to cultivate character and national awareness but also to enhance students' analytical, evaluative, and creative skills in comprehending and practising Pancasila's values (Sutrisno et al., 2023).

Higher Order Thinking Skills (HOTS) encompass the abilities to analyse (C4), evaluate (C5), and create (C6), as adapted in the Revised Bloom's Taxonomy (Anderson & Krathwohl, 2011). These skills are essential in the context of Pancasila Education, as students are expected to think critically about social issues, make decisions based on Pancasila values, and generate creative solutions for fostering a better national and civic life. In instructional practice, HOTS are implemented through various strategies such as problem-based learning, project-based learning, and inquiry learning (Asok et al., 2017; Vidergor & Krupnik-Gottlieb, 2015; Vu, 2025). However, the successful development of these skills depends not only on the instructional methods employed but also on the availability of assessment instruments that can measure higher-order thinking in a valid, reliable, and practical manner (Andayani & Madani, 2023).

HOTS-based assessment goes beyond testing students' ability to recall information; it evaluates how learners process, apply, and develop information within real-world contexts (Atmaja et al., 2024; Kurniawan & Utaminingsih, 2021). Therefore, it is crucial for teachers to have assessment instruments specifically designed to measure HOTS, whether in the form of essay tests, projects, case studies, or other types of authentic assessment (Brookhart, 2010).

Although theoretically important, the development and use of HOTS assessment instruments in Pancasila Education continue to face significant challenges in practice (Suyato et al., 2023). Many teachers struggle to construct HOTS-based questions, particularly in terms of formulating appropriate prompts, developing clear indicators, and designing effective scoring rubrics (Alotaibi, 2024; Sekwena, 2023).

There are a number of problems with adding HOTS to the Indonesian school system, especially in social studies. *First,* teacher readiness and skill. A lot of teachers know about HOTS, but they have trouble using it. Only 71% of teachers use HOTS in their lessons, and less than half of them plan lessons in a way that uses these skills (Sucipto et al., 2025). This indicates a significant disparity in the level of training and readiness among teachers. Teachers often don't have the skills they need to make good HOTS-based tests, which means that the goals of the curriculum and how it is taught don't match up (Kosasih et al., 2022; Zana et al., 2024)

Second, curriculum and Assessment Alignment The current curriculum puts a lot of emphasis on HOTS, but tests often focus on Lower-Order Thinking Skills (LOTS). This lack of alignment makes it hard to accurately assess students' higher-order cognitive skills. (Yuliana et al., 2024; Zana et al., 2024). In social studies, traditional methods put more value on facts than on HOTS skills like critical thinking and problem-solving (Supriatna et al., 2025).

Third, limitations on resources and support. Teachers need better teaching materials and support systems to help them use HOTS in their lessons more effectively (Faizah et al., 2021). Professional development programs are very important for giving teachers the skills they need to create and use HOTS-based assessments and lessons (Faizah et al., 2021; Zana et al., 2024). Students often don't have good critical thinking skills because they don't get enough practice with HOTS-based questions and problem-solving activities (Kosasih et al., 2022; Sanusi et al., 2025). Game-Based Learning (GBL) and the use of digital media are two new ways of teaching that have shown promise in improving students' critical thinking and engagement, but they need more development and support before they can be used by everyone(Ismail et al., 2024; Kosasih et al., 2022).

Fourth, differences by gender and grade level. Research indicates that gender may influence students' HOTS performance, with boys generally outperforming girls in higher-order thinking tasks. However, grade level appears to have minimal impact on HOTS development (Rifayanti & Mariana, 2024). Moreover, researchers have found that many test items claimed to assess HOTS still predominantly target lower-order cognitive skills, particularly remembering (C1) and understanding (C2) (Puspitasari, 2021). This issue is further exacerbated by the lack of training and technical support for teachers in developing HOTS-oriented assessment instruments at the school level.

Another challenge lies in the lack of habituation, time constraints, and the heavy administrative workload faced by teachers (Lidiawati & Aurelia, 2023). These factors often lead teachers to prefer simpler, quickly constructed assessment instruments over those requiring the design of higher-order thinking tasks. Additionally, some teachers perceive HOTS-based instruments as too complex to implement in large classrooms with a high number of students (Afifah & Retnawati, 2019).

Ideally, every assessment instrument should meet three essential criteria: validity, reliability, and practicality (Brookhart, 2010). Validity refers to the extent to which the instrument aligns with the intended learning objectives, reliability concerns the consistency of measurement results, while practicality relates to the ease with which the instrument can be used in real-world classroom settings. However, in actual practice, most existing studies tend to focus on issues of validity and reliability, often overlooking the aspect of practicality. Instruments that are valid and reliable, yet difficult to implement, may hinder the effectiveness of both teaching and assessment processes.

On the other hand, the emergence of the assessment for learning paradigm has emphasised that effective assessment should be an integral part of the learning process, rather than a mere administrative task (Black & Wiliam, 1998). Consequently, there is a need for HOTS assessment instruments that are not only substantively sound but also practical for teachers to use across various classroom contexts.

The Revised Bloom's Taxonomy serves as the primary foundation for developing HOTS indicators (Anderson & Krathwohl, 2011), while the 4Cs framework—Critical Thinking, Communication, Collaboration, and Creativity—represents the essential skill set that education must cultivate in the 21st century (Skills, 2009).

In the context of usability, a system or product is considered usable if it is easy to learn, efficient to use, easy to remember, error-tolerant, and provides user satisfaction (Nielsen, 1993). One of the simplest yet effective tools for measuring usability is the System Usability Scale (SUS), developed by Brooke (1996). SUS has been widely applied to evaluate the practicality of technological systems and software applications and has more recently been adopted in educational research to assess the usability of instructional instruments (Bangor et al., 2008).

This study introduces a novel approach by employing the System Usability Scale (SUS) to measure the practicality of a HOTS assessment instrument within Pancasila Education. This approach allows the instrument to be evaluated not only in terms of content quality but also through users' perceptions of its ease of use, clarity, and overall usability. By adopting a usability-based perspective, this study aims to contribute meaningfully to the development of assessment instruments that are more applicable and responsive to the practical needs of educational settings.

The novelty of this study lies in the application of usability principles to the evaluation of educational assessment instruments—an approach that remains relatively uncommon in the Indonesian context. Another key contribution is the development of an evaluation method based on teachers' perceptions as primary users, thereby enhancing the likelihood of adoption and successful implementation in classroom practice.

The findings of this study are expected to benefit multiple stakeholders. For teachers, the results can inform the selection or development of more practical assessment tools that align with instructional needs. For curriculum developers, the findings offer insights for designing more usable teaching materials that support effective and efficient learning processes. Furthermore, for educational researchers, this study provides a foundation for expanding instrument evaluation methods, particularly through a usability-based approach, to produce more adaptive and high-quality assessment tools.

Based on the foregoing discussion, the primary objective of this study is to measure the practicality of a HOTS assessment instrument in Pancasila Education using the SUS approach. In addition, the study aims to identify the strengths and weaknesses of the instrument's usability based on teachers' perceptions as end users. Furthermore, the research seeks to provide recommendations for developing more practical and applicable HOTS assessment tools for classroom use.

This section has outlined the importance of fostering HOTS in Pancasila Education, the challenges associated with developing appropriate assessment instruments, the gap between ideal and actual classroom conditions, and the innovation offered through the adoption of SUS as a means of evaluating instrument practicality. The study is expected

to contribute meaningfully to the development of learning instruments that are effective, user-friendly, and aligned with the demands of 21st-century education.

Method

This study employs a descriptive quantitative approach aimed at measuring the practicality of a HOTS assessment instrument in Pancasila Education using the SUS. A quantitative approach was chosen to obtain objective data that can be analysed statistically (Creswell, 2014). This design allows the researcher to capture measurable indicators of usability as perceived by teachers, enabling a structured evaluation of the instrument's effectiveness in classroom settings. The use of SUS as a standardised tool also facilitates comparison with broader usability benchmarks across educational contexts.

The participants in this study were fifth-grade primary school teachers in Bandung City who teach Pancasila Education and have implemented HOTS-based assessment. A total of 30 teachers were selected through purposive sampling, with the criteria of having a minimum of five years of teaching experience and having completed training on a 21st-century skills-based curriculum. This sample size meets the minimum requirement for instrument analysis using the SUS (Welda et al., 2020).

The research procedure was carried out in several stages. *First*, a HOTS-based assessment instrument was developed by referring to higher-order thinking indicators—namely C4 (analysing), C5 (evaluating), and C6 (creating)—as outlined in the Revised Bloom's Taxonomy (Anderson & Krathwohl, 2011). Each item in the instrument was designed to reflect domains of logical reasoning, decision-making, problem-solving, creativity, and creative thinking (Brookhart, 2010).

Second, the developed instrument underwent content validation by Pancasila Education experts to ensure the appropriateness of the content and alignment with learning indicators, and by assessment experts to verify that the items accurately represented the HOTS indicators and possessed sufficient content validity.

Third, the validated instrument was tested for practicality by teachers, who completed the SUS questionnaire after using the instrument in real classroom settings. *Finally,* the data from the SUS responses were analysed to determine the level of practicality of the developed assessment instrument.

The research instruments consisted of two main components. *First,* the HOTS assessment instrument, which took the form of higher-order essay questions, was developed based on the cognitive levels of analysing (C4), evaluating (C5), and creating (C6) as defined in the Revised Bloom's Taxonomy. The items were oriented toward higher-order thinking domains, including logical reasoning, decision-making, problem-solving, creativity, and creative thinking.

Second, the study employed the SUS, a ten-item questionnaire using a five-point Likert scale (Brooke, 1996). The SUS questionnaire was used to measure teachers' perceptions of the practicality of using the assessment instrument in everyday classroom settings. Data for this study were collected through two primary methods. First, the practicality scores were obtained from the SUS questionnaire completed by teachers after using the assessment instrument during instruction. Second, limited classroom observations were conducted to document the implementation of the instrument, serving as supporting data for triangulation to strengthen the findings derived from the questionnaire.

The quantitative data in this study were analysed through several steps. First, the

total SUS score for each respondent was calculated using the formula developed by Brooke (1996). Second, the scores were interpreted based on predefined usability categories: unacceptable (<50), marginal (50–69), and acceptable (>70). Third, the results were presented through frequency distributions and descriptive statistics, including the mean and standard deviation (Bangor et al., 2008).

Finding and Discussion

The practicality of the HOTS assessment instrument in Pancasila Education was measured through the completion of the SUS questionnaire by 30 teachers. Following the calculation, the average SUS score was 76.5, with a standard deviation of 8.2. The distribution of practicality categories based on SUS scores is presented in Table 1.

Table 1. Distribution of Instrument Practicality Categories

Practicality Category	SUS Score Range	Number of Teachers	Percentage (%)
Highly Practical	80–100	12	40%
Practical	70–79	15	50%
Moderately Practical	50–69	3	10%
Not Practical	<50	0	0%

Most respondents—90% of participating teachers—rated the instrument as falling within the practical to highly practical categories. Notably, none of the participants rated the instrument as not practical.

The SUS questionnaire consisted of 10 statement items. The average score for each item is presented in Table 2.

Table 2. Mean Score per Item of the SUS Questionnaire

No.	Statement	Mean Score
1	I think this instrument would be used frequently in teaching.	4.2
2	I found the instrument to be free of unnecessary complexity.	4.4
3	I found the instrument easy to use.	4.5
4	I think I would need technical support to use this instrument. (reversed)	1.9
5	I found the functions of the instrument to be well integrated.	4.3
6	I found inconsistency in the use of the instrument. (reversed)	2.0
7	I felt confident understanding how to use the instrument.	4.5
8	I found the instrument unnecessarily cumbersome. (reversed)	1.8
9	I felt confident using the instrument.	4.6
10	I had to learn a lot of things before I could use the instrument. (reversed)	2.1

According to the findings of the study, the average score on the SUS is 76.5. This score reflects the good opinions that teachers have regarding the simplicity of use, clarity of features, and effectiveness of the functions of the HOTS assessment instrument that was designed. According to the SUS interpretation criteria, this score is classified as "Good" and "Acceptable," and it is even getting close to the threshold of "Excellent"(Bangor et al., 2008). To calculate a legitimate final SUS score, the scores of negative items (items 4, 6, 8, and 10) have been reversed following the techniques that are considered conventional.

Based on these findings, it can be inferred that educators find the instrument to be not only practical but also not burdensome to use, and that it facilitates the application of higher-order thinking tests in the classroom. This instrument has met the practicality

indicators, as explained by van den Akker (1999), which states that an instrument is considered practical if it can be used in real-world conditions without requiring complex additional training. The high scores on the aspects of ease of use, clarity of instructions, and user confidence provide support for this assertion.

According to Welda et al. (2020), this instrument is also in accordance with usability standards, which describe it as being simple to learn, effective, simple to remember, and enjoyable for users. Based on these findings, it appears that the acceptance of the new instrument is facilitated by the good view that teachers have regarding the usefulness of the instrument. The perception of ease of use becomes a significant predictor of the acceptance of new technology or gadgets when considered in the context of the Technology Acceptance Model (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (Venkatesh et al., 2003). Both models were developed by Davis. As a result, high scores on this scale provide evidence that the possibility of continued usage in learning practices could be supported.

According to Trilling and Fadel (2009), the existence of practical instruments is of utmost importance to provide support for the implementation of the curriculum for the 21st century. This curriculum requires the development of HOTS, which includes critical thinking, problem-solving, and creative thinking. According to Brookhart (2010), the assessment of HOTS needs to be carried out with instruments that are not only valid in terms of their content but also efficient but also comfortable for teachers to use. A technique that is focused on usability, such as SUS, is therefore considered an innovation when it comes to evaluating the practicability of evaluation tools.

The HOTS assessment instrument that was developed as part of this research also addresses field concerns. According to Anggraini et al. (2019), HOTS-based evaluations in elementary schools are frequently seen as being difficult to implement and complicated. Although this instrument is designed to assess higher-order thinking skills, the research findings indicate that its use is not only straightforward and effective but also does not place an undue burden on educators (Epinasti et al., 2021). The fact that this is the case demonstrates that a successful HOTS assessment instrument does not have to be complicated; rather, it may be constructed to be effectively utilised in the classroom.

As an additional point of interest, the use of SUS in this setting illustrates an expansion of the capability of evaluation tools, which are often utilised in the process of measuring technological systems or software, into efficient methods for evaluating educational instruments that are not digital in nature. This represents the beginning of the integration between usability theory and the practice of building educational assessment systems specifically tailored to a given setting.

This instrument makes significant contributions within the framework of the Pancasila Education framework. Not only does assessment that is oriented towards HOTS enable teachers to measure students' factual mastery, but it also enables them to evaluate students' critical and reflective thinking skills, which are essential in the process of character development and the construction of national attitudes. Because this instrument is practical, it offers the possibility of developing a more comprehensive and pertinent evaluation of the Pancasila values in the context of the 21st century.

Additionally, from a methodological standpoint, this research paves the way for the creation of a theory about the measurement of practicality in the field of education. Since the beginning of time, the evaluation of practicability has frequently concentrated

on technical and technological elements. However, the findings of this research demonstrate that usability principles can be effectively utilised in the evaluation of curriculum-based assessment instruments. Because of this, there are now chances to develop more systematic standards for evaluating the usefulness of instructional devices.

In conclusion, this research contributes in the form of a holistic HOTS assessment instrument development model. This model not only examines the cognitive ability of students, but it also considers non-cognitive qualities such as creativity, teamwork, and problem-solving. The implementation of the Merdeka Curriculum, which places an emphasis on learning that is adaptive, contextual, and student-centred, can be supported by this paradigm, which can also be adapted for use in the development of additional instruments.

Conclusion

Based on the findings, it can be concluded that the developed HOTS assessment instrument is considered practical, with an average SUS score of 76.5, classified as "Good" and "Acceptable" according to standard SUS interpretation criteria. The instrument's practicality is reflected in teachers' positive perceptions regarding ease of use, clarity of instructions, feature effectiveness, and confidence in its application. These results suggest that the instrument can be efficiently implemented in real classroom settings without requiring extensive additional training. Accordingly, the initial expectation outlined in the introduction that a HOTS assessment tool should be not only valid but also practical to support 21st-century skills has been successfully achieved.

References

Adler-Beléndez, D., Hoppenstedt, E., Husain, M., Chng, E., & Schneider, B. (2020). How are 21st-century skills captured in makerspaces?: A review of the literature. *ACM International Conference Proceeding Series*, 40–45. https://doi.org/10.1145/3386201.3386214

Afifah, I. R. N., & Retnawati, H. (2019). Is it difficult to teach higher order thinking skills? *Journal of Physics: Conference Series*, *1320*(1). https://doi.org/10.1088/1742-6596/1320/1/012098

Alotaibi, S. K. G. (2024). The role of Islamic education teachers in enhancing higher-order thinking: a systematic review. *British Journal of Religious Education*. https://doi.org/10.1080/01416200.2024.2436921

Andayani, T., & Madani, F. (2023). Peran penilaian pembelajaran dalam meningkatkan prestasi siswa di pendidikan dasar. *Jurnal Educatio FKIP UNMA, 9*(2), 924–930. https://doi.org/10.31949/educatio.v9i2.4402

Anderson, L. W., & Krathwohl, D. R. (2011). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Allyn & Bacon.

Anggraini, N. P., Budiyono, & Pratiwi, H. (2019). Analysis of higher order thinking skills students at junior high school in Surakarta. *Journal of Physics Conference Series, 1211*(1). https://doi.org/10.1088/1742-6596/1211/1/012077

Asok, D., Abirami, A. M., Angeline CV, N., & Lavanya, R. (2017). Active learning environment for achieving higher-order thinking skills in engineering education. *Proceedings - 2016 IEEE 4th International Conference on MOOCs, Innovation and Technology in Education, MITE 2016*, 47–53. https://doi.org/10.1109/MITE.2016.69

Astuti, A. P., Aziz, A., Sumarti, S. S., & Bharati, D. A. L. (2019). Preparing 21st century teachers: Implementation of 4C character's pre-service teacher through teaching practice. *Journal of Physics: Conference Series*, *1233*(1). https://doi.org/10.1088/1742-6596/1233/1/012109

Atmaja, F. L., Hidayati, K., & Fineldi, R. J. (2024). Assessment of higher order thinking skills: Perspective of mathematics teacher. *AIP Conference Proceedings*, *2622*(1). https://doi.org/10.1063/5.0133815

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Bhardwaj, V., Zhang, S., Tan, Y. Q., & Pandey, V. (2025). Redefining learning: student-centered strategies for academic and personal growth. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1518602

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In *Usability Evaluation In Industry* (pp. 207–212). CRC Press. https://doi.org/10.1201/9781498710411-35

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.

Creswell, J. W. (2014). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Epinasti, G., Slamet, S. T. Y., & Yamtinah, S. (2021). Teacher's perception towards assessing of Higher Order Thinking Skills (HOTS) in elementary schools. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3516875.3516980

Faizah, R., Taqwa, M. R. A., & Istiyono, E. (2021). Senior high school student's higher order thinking skills based on gender and grade. *Journal of Physics: Conference Series*, *1918*(2). https://doi.org/10.1088/1742-6596/1918/2/022031

Fullan, M., & Scott, G. (2014). Education plus, new pedagogies for deep learning. *Collaborative Impact SPC,(July)*, 3–9.

Ismail, R., Retnawati, H., & Imawan, O. R. (2024). Construct validity of mathematics high order thinking skills instrument with cultural context: Confirmatory factor analysis. *AIP Conference Proceedings*, *3150*(1). https://doi.org/10.1063/5.0228143

Kocak, O., Coban, M., Aydin, A., & Cakmak, N. (2021). The mediating role of critical thinking and cooperativity in the 21st century skills of higher education students. *Thinking Skills and Creativity*, *42*, 100967. https://doi.org/10.1016/j.tsc.2021.100967

Kosasih, A., Supriyadi, T., Firmansyah, M. I., & Rahminawati, N. (2022). Higher-order thinking skills in primary school: Teachers' perceptions of Islamic education. *Journal of Ethnic and Cultural Studies*, *9*(1), 56–76. https://doi.org/10.29333/ejecs/994

Kurniawan, T. T., & Utaminingsih, S. (2021). Analysis of 4C-Based HOTS assessment module on critical thinking ability. *Journal of Physics: Conference Series*, *1823*(1). https://doi.org/10.1088/1742-6596/1823/1/012101

Lidiawati, K. R., & Aurelia, T. (2023). Kemampuan berpikir kritis siswa di Indonesia: Rendah atau tinggi. *Buletin KPIN*, *9*(02). https://buletin.k-pin.org/index.php/arsip-artikel/1200-kemampuan-berpikir-kritis-siswa-di-indonesia-rendah-atau-tinggi

Mahmud, M. M., & Wong, S. F. (2022). Digital age: The importance of 21st century skills among the undergraduates. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.950553

Nielsen, J. (1993). *Usability engineering*. Academic Press.

Puspitasari. (2021). Analisis kualitas soal berbasis HOTS pada pembelajaran IPS. *Jurnal Pendidikan Sosial Indonesia*, *3(1)*, 42–55.

Rifayanti, Z. E. T., & Mariana, N. (2024). Enhancing critical thinking and problem-solving skills in upper elementary students through game-based learning. *Perspektivy Nauki i Obrazovania*, *70*(4), 396–420. https://doi.org/10.32744/pse.2024.4.25

Sanusi, S., Suraiya, N., Nurasiah, N., Qismullah Yusuf, Y., Ridha, S., Hikmah, N., & Leni, L. (2025). Social knowledge management to optimize adolescents' social studies learning in Banda Aceh, Indonesia. *Cogent Education*, *12*(1). https://doi.org/10.1080/2331186X.2025.2506062

Sekwena, G. L. (2023). Active learning pedagogy for enriching economics students' higher order thinking skills. *International Journal of Learning, Teaching and Educational Research*, *22*(3), 241–255. https://doi.org/10.26803/ijlter.22.3.15

Shields, R., & Chugh, R. (2018). Preparing Australian high school learners with 21st century skills. *Proceedings of 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2018*, 1101–1106. https://doi.org/10.1109/TALE.2018.8615207

Skills, P. for 21st C. (2009). *Framework for 21st century learning.*

Sucipto, T. L. A., Sajidan, S., Akhyar, M., & Roemintoyo, R. (2025). Mapping teacher's readiness to improve the government policy in implementing higher-order thinking skills-based learning. *Qubahan Academic Journal*, *5*(1), 764–781. https://doi.org/10.48161/qaj.v5n1a1577

Supriatna, N., Winarti, M., & Wiyanarti, E. (2025). The effectiveness of using flipbooks as an interactive medium in social studies learning based on local wisdom to enhance critical thinking skills. *Review of Integrative Business and Economics Research*, *14*(2), 603–618. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85219026979&partnerID=40&md5=f2f1c6a3fafa55fd715ea543dedb090f

Sutrisno, S., Ponorogo, U. M., & Jumadi, I. J. (2023). Global citizenship education in the perspective of strengthening Pancasila student profiles. *Jurnal Civics: Media Kajian Kewarganegaraan*, *20*(2), 322–332. https://doi.org/10.21831/jc.v20i2.63983

Suyato, S., Hidayah, Y., Arpannudin, I., & Septiningrum, L. (2023). Revitalisasi pendidikan kewarganegaraan abad 21: Analisis keterampilan abad 21. *SOCIA: Jurnal Ilmu-Ilmu Sosial*, *19*(2), 78–84. https://doi.org/10.21831/socia.v19i2.60152

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Jossey-Bass.

van den Akker, J. (1999). Principles and methods of development research. In *Design Approaches and Tools in Education and Training* (pp. 1–14). Springer Netherlands. https://doi.org/10.1007/978-94-011-4255-7_1

van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. *SAGE Open*, *10*(1). https://doi.org/10.1177/2158244019900176

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425–478. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3375136

Vidergor, H. E., & Krupnik-Gottlieb, M. (2015). High order thinking, problem based and project based learning in blended learning environments. In *Applied practice for educators of gifted and able learners* (pp. 217–232). https://doi.org/10.1007/978-94-6300-004-8_11

Vu, X. H. (2025). Leveraging technology for higher-order thinking development: Instructional design strategies in higher education. *International Journal of Technologies in Learning*, *32*(2), 25–46. https://doi.org/10.18848/2327-0144/CGP/v32i02/25-46

Welda, W., Putra, D. M. D. U., & Dirgayusari, A. M. (2020). Usability testing website dengan menggunakan metode System Usability Scale (SUS)s. *International Journal of Natural Science and Engineering*, *4*(3), 152–161. https://doi.org/10.23887/ijnse.v4i2.28864

Yuliana, I. F., Faizah, F., Fatayah, F., & Priyasmika, R. (2024). Students' critical thinking skills using HOTS-based assessment on pH calculation. *AIP Conference Proceedings*, *3106*(1). https://doi.org/10.1063/5.0215096

Zana, F. M., Sa'dijah, C., Anwar, L., & Zulnaidi, H. (2024). Curriculum and teacher assessment practices in mathematics learning: Alignment with higher order thinking skills in Indonesian secondary schools. *Journal on Mathematics Education*, *15*(4), 1311–1334. https://doi.org/10.22342/jme.v15i4.pp1311-1334