Education Studies and Teaching Journal (EDUTECH)
Vol 2 (4) 2025 : 580-586

ANALYSIS OF THE EFFECTIVENESS OF PERFORMANCE ASSESSMENT TOOLS FOR DEVELOPING SCIENCE PROCESS SKILLS IN THE GENERAL NATURAL SCIENCES COURSE

ANALISIS EFEKTIVITAS ALAT PENILAIAN KINERJA DALAM MENGEMBANGKAN KETERAMPILAN PROSES SAINS PADA MATA KULIAH ILMU PENGETAHUAN ALAM UMUM

Rina Novianty
Institut Agama Islam Negeri Bone
*rinanovianty87@gmail.com

**Corresponding Author*

ABSTRACT
This study aims to analyze the effectiveness of performance assessment tools in developing students' science process skills (SPS) in the General Natural Science course. A total of 97 students from various study programs participated as respondents by completing a Likert-scale questionnaire consisting of 40 items. The findings indicate that the performance assessment tools fall into the effective category, with an average score of 4.154 (83.07%). All measured dimensions, conceptual understanding, learning objectives, thinking skills, contextual factors, and SPS were categorized as effective with a high level of consistency. These results align with the principles of authentic assessment and inquiry-based learning, which emphasize the role of process-oriented assessment in developing students' scientific competencies. This study contributes theoretically to strengthening the implementation of performance assessment in higher education and provides practical recommendations for lecturers in designing comprehensive assessment strategies.
Keywords: performance assessment, science process skills, inquiry learning

ABSTRAK
Penelitian ini bertujuan untuk menganalisis efektivitas alat penilaian kinerja dalam mengembangkan keterampilan proses sains (KPS) mahasiswa pada mata kuliah Ilmu Pengetahuan Alam Umum. Sebanyak 97 mahasiswa dari berbagai program studi berpartisipasi sebagai responden dengan mengisi kuesioner skala Likert yang terdiri dari 40 butir pernyataan. Hasil penelitian menunjukkan bahwa alat penilaian kinerja berada pada kategori efektif, dengan skor rata-rata sebesar 4,154 (83,07%). Seluruh dimensi yang diukur, yaitu pemahaman konseptual, tujuan pembelajaran, keterampilan berpikir, faktor kontekstual, dan KPS, dikategorikan efektif dengan tingkat konsistensi yang tinggi. Temuan ini sejalan dengan prinsip penilaian autentik dan pembelajaran berbasis inkuiri yang menekankan peran penilaian berorientasi proses dalam mengembangkan kompetensi ilmiah mahasiswa. Penelitian ini memberikan kontribusi teoretis dalam memperkuat implementasi penilaian kinerja di pendidikan tinggi serta memberikan rekomendasi praktis bagi dosen dalam merancang strategi penilaian yang komprehensif.
Kata kunci: penilaian kinerja, keterampilan proses sains, pembelajaran inkuiri

1. INTRODUCTION

Science process skills (SPS) are fundamental components in higher education that play an essential role in fostering students' critical thinking, analytical skills, and problem-solving abilities. SPS allows students not only to understand scientific concepts but also to apply them in real-world contexts through activities such as observation, classification, data interpretation, and drawing scientific conclusions. In the General Science course, mastery of SPS constitutes the foundation for developing scientific literacy and preparing students for advanced learning as well as the increasingly competitive demands of the workforce.

EDUTECH, 2 (4) 2025: 580-586, https://journal.ppipbr.com/index.php/edutech/index | DOI: https://doi.org/10.62207/qgqhkv23

However, the development of SPS often encounters challenges, primarily due to traditional learning approaches, the limited use of authentic assessments, and insufficient laboratory facilities. Conventional tests are widely considered inadequate to comprehensively measure students' scientific thinking processes. In this context, performance assessment emerges as a relevant evaluative approach to bridge this gap. It enables students to demonstrate competencies through real tasks such as experiments, data analysis, and presentation of investigative findings.

Several studies have shown that performance-based assessment significantly contributes to improved conceptual understanding and SPS (Srirahayu & Sulistyo Arty, 2023; Widiana, 2023). Nevertheless, most previous studies have focused on instrument development or its implementation at the primary and secondary education levels. Empirical studies evaluating the effectiveness of performance assessment tools within higher education, particularly in General Science learning, remain limited in both implementation and contextual factors influencing their success (Ovilia et al., 2024).

This research gap is reinforced by the bibliometric study of Sudirman et al. (2023), which concludes that investigations into performance assessment in Indonesian science education have yet to thoroughly explore its effectiveness or the contribution of each dimension to students' SPS development. According to Kurniawati (2021), assessment must encompass the entirety of scientific thinking processes to support meaningful and sustainable learning.

Based on this background, the objectives of this study are to:
(1) analyze the effectiveness of performance assessment tools in developing students' SPS, and
(2) identify the factors influencing the success of their implementation.
This study is grounded in the theories of authentic assessment and inquiry-based learning, which view assessment as an integral part of the learning process rather than merely a tool for evaluating outcomes.

2. METHODS

This research employed a quantitative descriptive design to provide a clear and measurable overview of the effectiveness of performance assessment tools in fostering students' SPS. The study population consisted of all first-semester students of the Faculty of Islamic Economics and Business (FEBI) at IAIN Bone in the 2024 academic year. First-year students were selected because they represent an appropriate group for evaluating the initial effectiveness of performance assessment tools in General Science learning.

A simple random sampling technique was used, resulting in a total sample of 97 students. This sampling method ensures equal opportunity for each member of the population to be selected, thereby minimizing bias and enhancing the validity of the findings. The research instrument was a Likert-scale questionnaire with five response categories: strongly disagree, disagree, neutral, agree, and strongly agree. This scale facilitates respondents' expression of their perceptions regarding the performance assessment tools used in learning.

Data were analyzed using quantitative descriptive techniques. Responses collected through Google Forms were processed using statistical software to measure the effectiveness level of the assessment tools. The analysis included percentages, means, and standard deviations for each item. The interpretation of effectiveness followed a Likert-based criterion (≥80% = effective).

3. RESULT AND DISCUSSION

3.1. Result

a. Respondent Profile

The study involved 97 students enrolled in the General Science course from various study programs, including Islamic Economics, Sharia Accounting, and Sharia Banking. These first-semester students had direct experience using performance assessment tools in the environmental pollution topic.

b. Overall Respon Distribution

From a total of 3,880 responses, the distribution is shown in Table 1:

Tabel 1 Response Distribution

Category	**Frequency**	**Percentage (%)**
Strongly Agree	1.212	31.24
Agree	2.022	52.11
Neutral	622	16.03
Dusagree	22	0.57
Strongly Disagree	2	0.05
Total	**3.880**	**100.00**

A total of **83.35%** of responses were positive, indicating a high level of acceptance of the performance assessment tools. Only **0.62%** were negative.

c. Analysis by Dimension

The instrument consisted of five main dimensions. Results are shown in Table 2.

Tabel 2. Dimensional Analysis

Dimension	**Items**	**Mean**	**%**	**SD**	Category
Conceptual Understanding & Tool Effectiveness	7	4.178	83.56	0.375	Effective
Learning Objectives & Engagement	7	4.130	82.59	0.407	Effective
Thinking & Analytical Skills	6	4.062	81.24	0.425	Effective
External & Contextual Factors	14	4.216	84.32	0.376	Effective
Science Process Skills	6	4.100	81.99	0.439	Effective
Overall Average	**40**	**4.154**	**83.07**	**0.404**	**Effective**

All dimensions scored within the effective category, with the highest score for external and contextual factors (84.32%).

3.2. DISCUSSION

a. Overall Effectiveness of Performance Assessment Tools

The overall score of 4.154 (83.07%) indicates that the assessment tools effectively supported the development of students' SPS. This aligns with Harlen's (2013) argument that performance assessment demonstrates higher construct validity in measuring scientific competencies compared to conventional tests.

The distribution of responses shows that 83.35% of students expressed positive perceptions (Agree and Strongly Agree), while only 0.62% gave negative responses. This distribution pattern indicates a very high level of acceptance toward the implementation of the performance assessment tools. However, it is important to note that 16.03% of students provided neutral responses, which may be interpreted as a group that has not yet experienced the optimal impact of the assessment tools. This group becomes a priority target for future instructional interventions and improvements

b. Strengths of the Implementation of the Performance Assessment Tools

The analysis of the top 10 items with the highest scores identified several key strengths in the implementation of the performance assessment tools. First, collaborative learning through student cooperation (item 28: 4.433) and group discussions (item 30: 4.433) proved to be highly effective. This aligns with social constructivist theory, which emphasizes the importance of social interaction in learning. When students work in groups to analyze cases of environmental pollution, they not only share knowledge but also construct a shared understanding through negotiated meaning and scientific argumentation.

Second, the quality of lecturer instruction emerged as a determining factor, reflected in the high scores for item 25 on lecturer explanations (4.392) and item 34 on lecturer engagement in providing feedback (4.392). This supports the findings of Hattie (2009), whose meta-analysis of more than 800 studies demonstrated that high-quality instructor feedback is one of the factors with the largest effect size on student achievement.

Third, the contextualization of learning through the use of real examples (item 26: 4.278) was shown to be effective in enhancing the relevance and meaningfulness of learning. When students study environmental pollution through local cases familiar to them such as river contamination or air pollution in their own city the learning experience becomes more meaningful, thereby increasing intrinsic motivation.

Fourth, the performance assessment tools were shown to be effective in facilitating conceptual understanding (item 1: 4.351) and students' ability to draw conclusions from experiments (item 37: 4.299). This indicates that the hands-on and minds-on approaches embedded in performance assessment are capable of developing students' higher-order thinking skills.

c. Areas Requiring Improvement

Although the overall findings indicate effectiveness, the analysis of the bottom 10 items reveals several aspects that require improvement. First, the ability to formulate hypotheses (item 18: 3.763) emerged as the most significant weakness. Hypothesizing skills are a crucial component of scientific inquiry, requiring deductive thinking and reasoning abilities. Lawson (1995) explains that these skills develop through cognitive stages that necessitate explicit practice and appropriate scaffolding. To address this issue, a gradual approach is needed, beginning with guided inquiry and progressing toward open inquiry, accompanied by explicit instruction on the structure and characteristics of a sound scientific hypothesis.

Second, students' confidence (item 7: 3.938) still needs to be strengthened. Self-efficacy is an important predictor of persistence and achievement in science learning (Bandura, 1997). The relatively low score in this aspect can be improved through mastery experiences (providing gradual success experiences), vicarious experiences (peer modeling), verbal persuasion (positive encouragement), and managing emotional states (reducing anxiety during science learning).

Third, the ability to comprehensively measure science process skills (item 13: 3.876) indicates the need for refinement in the design of the assessment instrument. Rustaman et al. (2005) identify 13 components of SPS that must be assessed holistically, including observation, classification, measurement, communication, prediction, inference, variable identification, data

tabulation, graphing, describing variable relationships, hypothesis formulation, experimental design, and conducting experiments. The assessment instrument must therefore be redesigned with a more comprehensive rubric that measures all of these components.

Fourth, the aspects of technology use (item 23: 4.072) and resource availability (item 21: 4.041), although still within the good category, show room for improvement. In the digital era, the integration of technology in science learning is not optional but essential. The use of virtual laboratories, simulations, and data loggers can enhance accessibility and engagement in science learning, especially for topics that are challenging to observe directly, such as microplastic pollution or climate change.

.

d. Theoretical Implications

The findings of this study provide important contributions to several theories of learning and assessment. First, the study reinforces the theory of authentic assessment proposed by Wiggins (1990) and Mueller (2005), which emphasizes the importance of assessments that reflect the complexity of real-world tasks and competencies. Performance assessment within the context of environmental pollution enables students to demonstrate their understanding and skills in relevant and meaningful situations, rather than merely recalling facts on written tests.

Second, the effectiveness of collaborative learning identified in this study supports the theory of situated cognition (Lave & Wenger, 1991), which posits that learning occurs most effectively within authentic social and cultural contexts. Students not only learn about environmental pollution but also learn through participation in a learning community that shapes their identities as prospective scientists or environmental practitioners.

Third, the findings regarding the importance of formative feedback support the theory of formative assessment developed by Sadler (1989) and Black & Wiliam (1998). Effective feedback not only informs students about what is incorrect but also explains why it is incorrect, what constitutes correct understanding, and how they can improve. This fosters a cycle of continuous improvement in the learning process.

3.3. Practical Implications for Learning

Based on the research findings, several practical implications can be formulated.

3.3.1. Curriculum Development:

Performance assessment tools need to be systematically integrated into curriculum design, not as an add-on but as an integral component of the learning process. This requires alignment among learning outcomes, learning activities, and assessment methods in accordance with the principles of constructive alignment (Biggs, 1996).

3.3.2. Lecturer Professional Development:

The study highlights the importance of instructional quality. Professional development programs should therefore focus on:

designing and implementing performance assessments, techniques for providing effective formative feedback, strategies for facilitating collaborative learning, and the use of scaffolding to develop higher-order thinking skills.

3.3.3. Instructional design:

Strengthening is needed in areas that received lower scores, particularly in developing students' ability to formulate hypotheses and their self-confidence.This can be achieved through: explicit instruction on the processes of scientific inquiry, gradual guided inquiry transitioning from structured to open inquiry, peer learning and collaborative problem solving, and assigning meaningful and achievable tasks to build self-efficacy.

3.3.4. Learning infrastructure:

Although technology and resource scores were already good, continuous investment remains necessary in: the development of laboratories and practical equipment, access to digital technologies and software for data analysis, the development of learning resources such as case studies and multimedia materials, and the creation of learning spaces conducive to collaborative learning.

4. CONCLUSION

Based on a comprehensive analysis of data from 97 respondents across 40 questionnaire items, this study yields several key conclusions. First, the performance assessment tools developed were proven effective in enhancing students' science process skills in the General Science course on environmental pollution, with an overall effectiveness score of 4.154 on a 5-point scale (83.07%). This level of effectiveness is supported by the high proportion of positive responses, which reached 83.35% of the 3,880 responses analyzed.

Second, all five measured dimensions were categorized as effective, with External and Contextual Factors achieving the highest score (4.216 or 84.32%), followed by Conceptual Understanding & Tool Effectiveness (4.178), Learning Objectives & Engagement (4.130), Science Process Skills (4.100), and Thinking & Analytical Skills (4.062). The relatively small variation in scores across dimensions (SD = 0.375–0.439) indicates consistent effectiveness of the assessment tools.

Third, the factors contributing most significantly to the tools' effectiveness include: collaborative learning through peer cooperation and group discussions, the quality of lecturer instruction in providing explanations and feedback, the use of real-world examples and contextualized learning, and the tools' ability to facilitate conceptual understanding and students' ability to draw conclusions.

Fourth, several aspects still require improvement, including: students' ability to formulate hypotheses, students' confidence in learning science, the comprehensiveness of the instrument in measuring all components of science process skills, and the optimization of technological integration and learning resource utilization.

Based on the findings and conclusions of this study, several strategic recommendations are proposed: For educational institutions: Adopt performance assessment tools as an integral component of the assessment system in higher education; Provide adequate infrastructural support for the implementation of performance-based assessments; Develop policies that encourage the use of authentic assessment in science learning; Allocate funding for the development of laboratories and learning technologies. For curriculum developers: Design curricula grounded in the principles of constructive alignment that integrate performance assessment tools; Develop learning outcomes that emphasize science process skills and higher-order thinking; Prepare contextual and relevant teaching materials and assessment tools aligned with contemporary environmental issues; Design a progression framework for developing scientific inquiry skills, moving from guided to open inquiry.

For lecturers/instructors: Enhance competencies in designing and implementing performance assessments through continuous professional development; Develop comprehensive assessment rubrics that measure all aspects of science process skills; Provide explicit instruction and appropriate scaffolding to improve students' ability to formulate hypotheses; Facilitate collaborative learning through structured cooperative learning strategies; Provide specific, constructive, and timely formative feedback; Apply mastery learning approaches to build students' confidence.

For future researchers: Conduct longitudinal studies to measure the long-term impact of performance assessments on students' scientific literacy; Develop experimental studies with control groups to validate the causal effectiveness of the assessment tools;

Undertake comparative studies of various performance assessment models to identify best practices; Explore the use of digital technologies such as virtual reality and artificial intelligence in science performance assessment; Investigate individual and contextual factors that mediate the effectiveness of performance assessment tools.

For students: Utilize performance assessment tools as opportunities to develop scientific skills and critical thinking; Actively participate in collaborative learning and group discussions; Use feedback from lecturers to support continuous improvement; Develop self-regulated learning strategies in science learning.

The implementation of these recommendations is expected to enhance the quality of science education and foster more comprehensive scientific literacy among students, not only in the General Science course but also across other science-related courses in higher education.

5. REFERENCES

Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman.

Biggs, J. (1996). Enhancing Teaching Through Constructive Alignment. Higher Education, 32(3), 347–364. https://doi.org/10.1007/BF00138871

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

Harlen, W. (2013). Assessment & inquiry-based science education: Issues in policy and practice. Global Network of Science Academies.

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Kurniawati, A. (2021). Science process skills and its implementation in the process of science learning evaluation in schools. Journal of Science Education Research, 5(2), 16–20. https://doi.org/10.21831/jser.v5i2.44269

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge University Press.

Lawson, A. E. (1995). Science teaching and the development of thinking. Wadsworth.

Mueller, J. (2005). The authentic assessment toolbox: Enhancing student learning through online faculty development. Journal of Online Learning and Teaching, 1(1), 1–7.

Ovilia, G., Sakina, N., Nurlatifah, N., & Nahadi, N. (2024). Analysis the implementation of science process skills (SPS) assessment in chemistry learning in Indonesia: A systematic literature review. Hydrogen: Jurnal Kependidikan Kimia, 12(6). https://doi.org/10.33394/hjkk.v12i6.13366

Rustaman, N. Y., Dirdjosoemarto, S., Yudianto, S. A., Achmad, Y., Subekti, R., Rochintaniawati, D., & Nurjhani, M. (2005). Strategi belajar mengajar biologi. Jurusan Pendidikan Biologi, FPMIPA UPI.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144. https://doi.org/10.1007/BF00117714

Srirahayu, R. R. Y., & Sulistyo Arty, I. (2023). Pengembangan instrumen experiment performance assessment untuk menilai keterampilan proses sains dan kerja sama. Jurnal Penelitian dan Evaluasi Pendidikan, 22(2). https://doi.org/10.21831/pep.v22i2.20270

Sudirman, S., Ramdani, A., Doyan, A., Arian Sani, Y., & Rokhmat, J. (2023). A bibliometric analysis performance assessment of science education on science process skills. Path of Science, 9(3–4). https://doi.org/10.22178/pos

Widiana, I. W. (2023). Performance and project assessment in science learning. Jurnal Pendidikan dan Pengajaran, 55(2). https://doi.org/10.23887/jpp.v55i2.48942

Wiggins, G. (1990). The case for authentic assessment. Practical Assessment, Research, and Evaluation, 2(2), 1–3. https://doi.org/10.7275/kvfd-xm56