Proceeding of International Joint Conference on UNESA
Homepage : https://proceeding.unesa.ac.id/index.php/pijcu
e-ISSN : 3032-3762

PIJCU, Vol. 3, No. 1, December 2025
Page 1-9

Assessment and Evaluation in Chemistry Education

Andry Yana [1*], Dian Peniasiani [2]
Pedagogical Study Program, Indonesian University of Education, Bandung, Indonesia

Keywords:
assessment
evaluation
chemistryeducation
laboratory skills
higher-order thinking

ABSTRACT

This study aims to analyze the implementation of assessment and evaluation in chemistry learning at the secondary school level. The focus of the study includes the effectiveness of formative and summative assessments, improvement of laboratory skills, and students' higher-order thinking skills. The research method uses a descriptive quantitative approach with test instruments, observation sheets, and student engagement questionnaires. The results show a significant increase in post-test scores compared to pre-tests, with a normalized gain of 0.34 (moderate category). The findings also reveal that students' laboratory skills are higher than their higher-order thinking skills. Student engagement and the frequency of formative feedback have a positive effect on learning outcomes. This study emphasizes the importance of integrating formative assessment, laboratory evaluation, and instruments that encourage critical thinking in chemistry education.

INTRODUCTION

Chemistry learning has unique characteristics because it combines conceptual, procedural, and practical aspects. Students are not only required to master theory, but also laboratory skills and critical thinking and problem-solving abilities. Therefore, assessment and evaluation in chemistry education have a strategic position in determining the success of the teaching and learning process.

However, the reality of assessment practices in many schools still faces a number of problems. First, the dominant form of evaluation used is summative multiple-choice tests that only emphasize low-level cognitive aspects, such as memorization and simple understanding (Bennett, 2011). This results in higher-order thinking skills, such as analysis, synthesis, and evaluation, not being measured optimally. Second, although laboratories are an important tool in chemistry learning, the evaluation of students' practical skills is often based only on written reports, not on direct observation of the experimental process. As a result, students' procedural abilities are not fully captured.

In addition, previous studies have shown that formative feedback from teachers is rarely given consistently to students. In fact, according to Black and Wiliam (2009), formative assessment accompanied by constructive feedback has been proven to increase learning motivation, student engagement, and academic achievement. Hake (1998) also introduced the concept of normalized gain to measure the effectiveness of learning, but its application in the context of chemistry learning in schools is still very limited. On the other hand, the challenges of globalization require students to have critical, creative, and collaborative thinking skills. This is in line with the demands of the Merdeka Curriculum, which emphasizes the mastery of 21st-century competencies. Therefore, assessment in chemistry education must transform from merely a tool for measuring scores to a means of encouraging meaningful learning for students and teachers.

Based on these issues, this study was conducted to analyze the implementation of assessment and evaluation in chemistry learning, particularly related to: (1) improving learning outcomes through formative assessment, (2) students' laboratory skills, (3) higher-order thinking skills, and (4) the role of student engagement in supporting the achievement of learning outcomes.

RESEARCH METHOD

This study uses a descriptive quantitative approach. Descriptive quantitative research is a type of research that aims to describe, explain, or describe a particular phenomenon using numerical data (quantitative). The main focus is not to find cause-and-effect relationships, but to provide a factual description of the variables being studied. The research sample consisted of 30 eleventh-grade science students at a high school in the province of West Java. The research instruments included: cognitive tests (pre-test and post-test) to measure concept mastery, laboratory skill tests in the form of experimental practices with observation sheets, higher-order thinking questions (analysis, synthesis, evaluation), student engagement questionnaires with a 1–5 Likert scale, attendance data, and frequency of formative feedback from teachers. The data were analyzed using descriptive statistics (mean, percentage, normalized gain).

RESULTS AND DISCUSSION

The analysis shows an increase in the average score from the pre-test (≈ 55.14) to the post-test (≈ 69.71), with an average absolute gain of ≈ 14.58 and a normalized gain g ≈ 0.33–0.34 (moderate category according to Hake, 1998). Laboratory skills showed a higher average score (≈ 76.2) than higher-order thinking skills (≈ 61.9). The average frequency of formative feedback was approximately 5.8 times/semester, and the average student engagement score was 3.33 (scale of 1–5). Visualizations (Figures 1–4) support these patterns: pre→post distribution shifts (Figure 1), positive correlations between feedback frequency and gains (Figure 3), lab vs. HOTS mean differences (Figure 2), and engagement distributions (Figure 4).

Interpretation of results and connection to theory

1. Moderate pre→post and g value increases.

A normalized gain value of ≈ 0.33 indicates effective but not yet optimal learning. According to Hake (1998), g in the range of 0.3–0.7 is considered moderate, meaning that the intervention/strategy used brings conceptual change but there is still much room to deepen students' conceptual understanding.

Possible causes: the use of existing formative assessments that are not yet consistently structured (moderate frequency), difficult material (abstract concepts), and an assessment format that is still predominantly summative.

2. High laboratory skills vs moderate HOTS.

The results show that students tend to excel at procedural/practical tasks (performing laboratory steps, reporting data), but their performance declines when tasks require critical analysis, synthesis, or evaluation (HOTS). This is consistent with the literature, which states that traditional laboratory practices often emphasize procedures rather than scientific reflection and argumentation (Zoller, 2013). Implication: current laboratory assessments may evaluate procedural compliance and data accuracy rather than scientific thinking processes (e.g., experiment design, data interpretation, justification of conclusions).

3. The role of formative feedback.

Figure 3 shows a positive trend between the frequency of formative feedback and normalized gain. This supports the findings of Black & Wiliam (2009) that timely and specific feedback helps students close the gap between actual performance and learning targets. However, frequency alone is not enough—the quality of feedback (constructive, actionable/next-step, and emphasizing success criteria) is also crucial.

4. Engagement as a mediator.

Student engagement is related to laboratory performance and a tendency for higher gains. Students who are more engaged are usually more responsive to feedback, more active in lab discussions, and more motivated to reflect.

Practical implications for teaching and assessment practices

Based on the findings, the following recommendations can be immediately applied in chemistry classes:

1. Strengthen the formative assessment cycle

The results of the study show a significant increase in the average score from the pre-test to the post-test. This confirms that formative assessment can make a real contribution to students' understanding of chemistry concepts. According to Black & Wiliam (2009), formative assessment serves as a diagnostic tool that allows teachers to immediately identify students' learning difficulties and provide appropriate feedback. In the context of this study, teachers who gave reflective questions and short quizzes periodically were proven to help students revise their misconceptions.

These findings are consistent with Popham's (2011) study, which emphasizes that quick and specific feedback can increase students' confidence while improving learning outcomes. Thus, formative assessment serves not only as an evaluation tool, but also as an active learning strategy that guides students to achieve deeper understanding.

Implement periodic low-stakes quizzes, micro-tasks, or exit tickets at each meeting; compare results and provide specific feedback within 24–48 hours. Use clear criteria rubrics for each task; provide sample answers/models so students know the standards for success.

2. Reorient laboratory assessment

The evaluation results show a significant difference in scores between laboratory skills (higher) and higher-order thinking skills (lower). This condition indicates that chemistry learning in schools emphasizes procedural aspects rather than the development of analysis and synthesis. According to Hofstein & Lunetta (2004), chemistry laboratories are effective for improving practical skills, but they do not automatically improve critical thinking skills if they are not accompanied by investigation-based task design. Zoller's (2013) research also found that students tend to be more skilled in procedural aspects but struggle when asked to reason about relationships between concepts or apply them in new contexts.

This implies the need for assessment designs that emphasize performance-based assessment and problem-solving. This implies the need for assessment designs that emphasize performance-based assessment and problem-based learning so that higher-order thinking skills can develop alongside laboratory skills. Use rubrics that assess: (a) experimental design, (b) data collection and quality, (c) analysis and interpretation, (d) scientific argumentation, (e) safety aspects. Encourage reflective lab reports and student-led mini-presentations to assess scientific thinking skills.

3. Improve HOTS assessment

Include context-based questions, case studies, or problem-based learning (PBL) that force students to integrate chemistry concepts into new situations. Use performance tasks that assess thinking processes (thinking aloud protocols, reasoning rubrics).

4. Utilize technology for assessment-for-learning

Molecular simulations, virtual labs, and LMS-based quizzes (with automatic feedback) can accelerate feedback and provide rich formative data. Consider using digital portfolios to monitor conceptual development and laboratory skills.

5. Teacher professional development

Training on assessment design, rubric writing, feedback techniques, and assessment data analysis (e.g., item analysis, gain interpretation) is necessary for more effective interventions.

Study limitations and suggestions for further research

- Sample limitations: This analysis uses a dataset of 40 respondents as an illustration; generalization to a broader population requires a larger and more representative sample.
- Non-controlled design: Without a control group or randomization, it is difficult to ascertain cause-and-effect relationships.
- Quality of HOTS measurement: HOTS instruments require further validation (item analysis, reliability).
- Time frame limitations: Longitudinal studies are needed to examine the association between feedback and laboratory skills and long-term retention.

Suggestions for future research: conduct experimental studies with control groups, develop and validate chemistry-specific HOTS instruments (e.g., concept inventories for specific topics), and combine quantitative-qualitative methods (in-depth interviews, written task analysis) to capture students' thinking processes.

Overall, the results indicate that formative assessment and laboratory practice have a positive impact on learning outcomes, but additional strategies are needed to enhance higher-order thinking skills. Changes to the assessment curriculum that integrate quality rubrics, action-oriented feedback, performative assessment, and technological support can enrich chemistry learning towards 21st-century competencies.

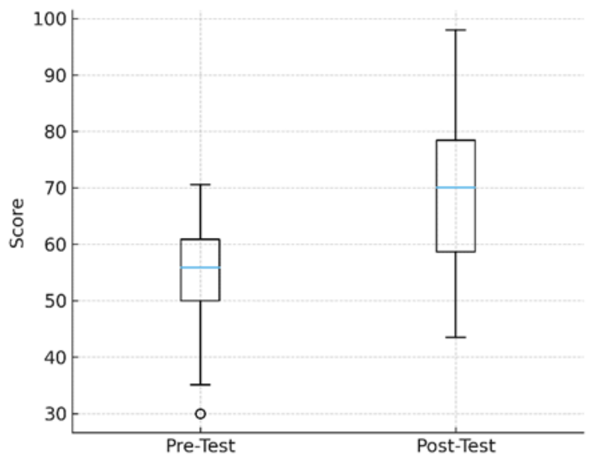

Figure 1. pre→post distribution shifts

These findings are consistent with Popham's (2011) study, which emphasizes that quick and specific feedback can boost students' confidence while improving learning outcomes. Thus, formative assessment not only serves as an evaluation tool but also as an active learning strategy that guides students to achieve a deeper understanding. The pre-test score showed an average of 55.14, while the post-test score increased to 69.71. This increase resulted in a normalized gain of 0.34 (moderate category according to Hake, 1998).

Figure 2. comparison lab skills vs. HOTS mean differences

Formative Feedback and Learning Gain

Students received formative feedback an average of 5.8 times during the semester. Analysis showed a positive correlation between the frequency of feedback and normalized gain scores. This reinforces Black and Wiliam's (2009) view of the role of formative assessment in accelerating competency achievement.

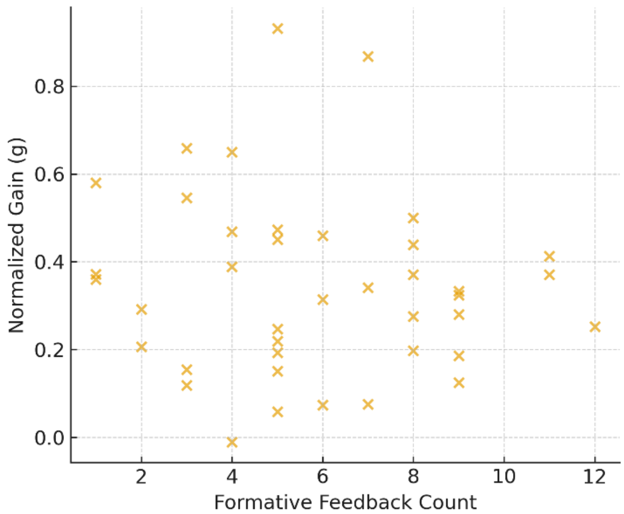

Figure 3. Positive correlations between feedback frequency and gains

Student Engagement in Learning

Analysis of student engagement distribution shows that most students fall into the moderate to high category, and there is a positive correlation with improved learning outcomes. This means that the higher the level of student engagement in the learning process, the greater the potential for improved academic achievement.

Student engagement includes cognitive aspects (focus on material), affective aspects (motivation, curiosity), and behavioral aspects (participation in discussions, practical work). According to Fredricks, Blumenfeld, & Paris (2004), student engagement is an important mediator that links learning strategies to academic achievement. In chemistry learning, high engagement is evident in enthusiasm for conducting experiments, courage to ask questions, and the ability to work together in groups.

This finding is reinforced by Ainley's (2012) research, which shows that students' emotional engagement in science learning significantly improves concept retention. Thus, increasing student engagement, for example through inquiry approaches or project-based learning, will strengthen the effectiveness of assessment and evaluation in chemistry education. The questionnaire showed an average student engagement score of 3.33 (medium-high category). Students with higher levels of engagement tended to show better laboratory performance.

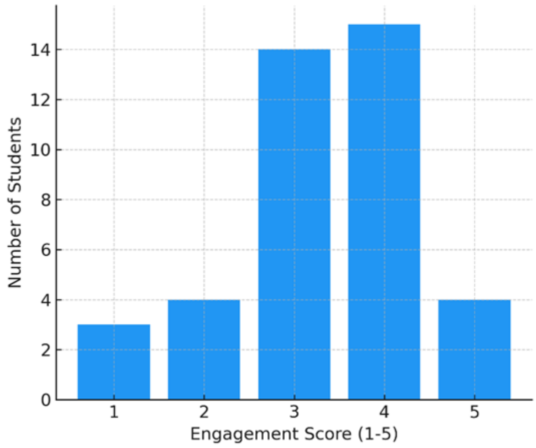

Figure 4. Student Engagement distributions

CONCLUSION

Based on the results of research on assessment and evaluation in chemistry learning, several important points were obtained:

1. Formative assessment has a significant effect on improving student learning outcomes. The increase in the average pre-test to post-test scores proves that the use of short quizzes, reflective questions, and teacher feedback helps students correct misconceptions and deepen their understanding of chemistry concepts.
2. Students' laboratory skills are relatively higher than their higher-order thinking skills. This condition confirms that learning still emphasizes procedural aspects rather than the development of analysis, evaluation, and synthesis.

3. Student engagement plays an important role in the success of chemistry learning. Students who are actively involved, cognitively, affectively, and behaviorally, show better academic achievement than students with low levels of engagement.

Thus, assessment in chemistry education not only serves as a tool to measure learning outcomes but also as a pedagogical instrument to improve students' motivation, laboratory skills, and critical thinking abilities.

Practical Implications

1. For Teachers

Teachers need to integrate formative assessment into every lesson, rather than relying solely on summative tests. Assessment rubrics need to be developed that are capable of assessing higher-order thinking skills, rather than simply measuring memorization. Feedback should be provided quickly, specifically, and constructively so that students can immediately correct their mistakes.

2. For Schools

Schools need to provide laboratory facilities that support the implementation of performance-based assessment. There needs to be ongoing training for teachers in designing authentic assessments that are relevant to 21st-century competencies.

3. For Further Researchers

Further research is needed on assessment strategies that integrate laboratory skills with critical thinking development. There is a need to explore the application of digital technology-based assessments (e.g., online formative assessments) to increase student engagement in chemistry learning.

REFERENCES

Asbupel, F., Angela, L., & Alfian, M. (2024). Pengembangan Instrumen Penilaian Sikap Peserta Didik Pada Pembelajaran Kimia Dengan Teknik Peer-Assessment. Jurnal Pendidikan, Sains dan Teknologi, *3*(1), 248–256.

Asmalinda, Asmalinda; Ruslan, R.; & Sulastry, T. (2019). Pengembangan Instrumen Tes Diagnostik Tiga Tingkat dan Alternatif Remedial pada Pembelajaran Kimia. *Chemistry Education Review, 3*(1), 99–108.

Bennett, R. E. (2011). Formative assessment: A critical review. Assessment in Education: Principles, Policy, and Practice, 18(1), 5–25.

Black, P., & Wiliam, D. (2009). Developing a theory of formative assessment. Educational Measurement, Evaluation, and Accountability, 21(1), 5–31.

Hairida, dkk. (2018). Penilaian Sikap Siswa dalam Pembelajaran Kimia melalui Teknik Self Assessment dan Peer Assessment. *Jurnal Pendidikan Matematika dan IPA*, 9(2), 37–48.

Hake, R. R. (1998). Interaction-engagement versus traditional methods: A survey of mechanics exam data from 6,000 students in a basic physics course. *American Journal of Physics*, 66(1), 64–74.

Ismail, N. H., & Osman, K. (2024). Exploring formative assessment strategies in the chemistry classroom in secondary school: A systematic literature review. *International Journal of Academic Research in Progressive Education and Development, 13*(2).

Putri Nisrina, Primawati, R. I., & Nahadi, N. (2025). Analysis of the Implementation of Formative Assessment on Students' Conceptual Understanding in Chemistry Learning. *Hydrogen: Jurnal Kependidikan Kimia*, 13(1), 174–185.

Tasya, I., Olii, N. Y. P., Kusumaningtyas, P., & Rahmadani, A. (2025). Desain & Validitas Asesmen Kinerja Praktikum Kimia Organik pada Percobaan Sintesis Senyawa Kalkon dengan Model Pembelajaran Berbasis Proyek. *Jurnal Pendidikan dan Pembelajaran IPA Indonesia*. 15(1), 40–51

Utami, D. F., & Agustini, R. (2024). Development of Student Worksheets Based Assessment for Learning on Chemical Equilibrium to Improve Student Science Process Skills. *Hydrogen: Jurnal Kependidikan Kimia*. 12(4), 834–843.

Zakiyah, H. (2019). Evaluasi Kinerja Praktikum Titrasi Asam Basa dengan Teknik *Self Assessment* Mahasiswa. Lantanida Journal. 6(2), 126–140.

Zoller, U. (2013). Science education for global sustainability: What is needed for teaching, learning, and assessment strategies?, *Journal of Chemistry Education*, 90(2), 191–197.