

Performance Profiling of Physics Students Across Bloom's Taxonomy Levels in a Malaysian Higher Education Institution

Siti Fatimah Saipuddin*, Wan Rozianoor Mohd Hassan

Faculty of Applied Sciences, University Technology MARA, 40450 Shah Alam, Selangor, Malaysia

*Corresponding author

DOI: <https://dx.doi.org/10.47772/IJRISS.2025.908000067>

Received: 24 July 2025; Accepted: 30 July 2025; Published: 28 August 2025

ABSTRACT

This study explores the academic performance of undergraduate Physics students in a Malaysian higher education institution through the lens of Bloom's Taxonomy. The primary objective was to assess how students performed on examination items aligned with different cognitive levels, specifically comparing lower-order thinking skills in CLO1 and higher-order thinking skills in CLO2. A mixed-methods approach utilizing a sequential explanatory design was employed. In the quantitative phase, exam scores from 59 students were analysed using descriptive statistics, paired samples t-tests, and Pearson correlation analysis. Results revealed that students scored significantly higher in CLO2 ($M = 27.78$) than in CLO1 ($M = 24.46$), with a strong positive correlation ($r = .75$) between the two. These findings suggest that students may be better equipped or more engaged when tackling analytical and evaluative tasks, challenging traditional assumptions about cognitive difficulty in assessments. The study also discusses potential instructional and assessment-related factors contributing to this trend. Based on the results, recommendations are proposed to improve assessment practices and cognitive alignment in Physics education. This research contributes to the growing discourse on outcome-based education, offering practical insights for curriculum designers, instructors, and policymakers.

Keywords— Bloom's Taxonomy, Physics Education, Higher-Order Thinking, Student Performance, Course Learning Outcomes (CLO), Assessment Design

INTRODUCTION

In the rapidly evolving world of higher education, the efficacy of evaluation methods plays a pivotal role in influencing students' cognitive development and scholarly achievement. A commonly utilized framework for the design and assessment of educational objectives is Bloom's Taxonomy, which organizes cognitive skills into a hierarchical structure from lower-order to higher-order thinking: Remember, Understand, Apply, Analyze, Evaluate, and Create (Anderson & Krathwohl, 2001). This taxonomy not only assists in developing complete learning outcomes, but it also encourages the alignment of teaching methods, assessment design, and expected student performance (Krathwohl, 2002; Zoller, 2018).

The significance of cognitive development in the field of physics cannot be overstated, as it facilitates a deeper understanding of concepts, enhances analytical thinking, and aids in the resolution of scientific challenges. Physics serves as a distinctive discipline for evaluating students' comprehension across various cognitive levels, as it necessitates both abstract reasoning and numerical application. Nonetheless, studies indicate that university physics assessments frequently prioritise lower-order cognitive abilities such as recall and understanding, while insufficiently addressing higher-order thinking skills like evaluation and creativity (Stelzer et al., 2020; Nieminen et al., 2021). The absence of alignment may lead students to engage in superficial learning, hindering their ability to employ deeper scientific reasoning, which is essential for advancing scientific inquiry and generating innovative concepts (Redish, 2003; Caballero & Wilcox, 2017).

The Malaysian Qualifications Agency (MQA) demands the implementation of Outcome-Based Education (OBE) frameworks within higher education institutions. The frameworks exhibit a strong connection to

Bloom's Taxonomy. This aims to ensure that the learning outcomes of the course align with national quality standards and meet global competency expectations (MQA, 2020). Despite the ongoing nature of this project, there is a notable lack of empirical studies examining the distribution of physics test questions across various cognitive levels and the subsequent impact on student performance. To enhance instructional strategies, increase the validity of evaluations, and encourage deeper thinking among physics students, it is crucial to comprehend this relationship.

This investigation aims to address that gap by examining the performance of undergraduate physics students on test questions categorised at various levels of Bloom's Taxonomy. The aim of this study is to provide educators, curriculum designers, and decision-makers with valuable insights on enhancing the alignment of assessment strategies with the desired learning objectives in physics education by examining student performance across various cognitive domains.

LITERATURE REVIEW

Bloom's Taxonomy remains a fundamental concept in education, particularly regarding the development of assessments designed to evaluate students' cognitive engagement. The classification system, initially developed by Bloom et al. in 1956 and subsequently revised by Anderson and Krathwohl in 2001, organises educational objectives into a structured hierarchy, progressing from lower order thinking skills to higher order thinking skills, including analysis, evaluation, and creation. These classifications play a significant role in the development of curricula and the planning of assessments within higher education.

This perspective is crucial in every aspect of science education, particularly in the field of physics. Students must engage in deeper understanding beyond merely recalling scientific facts. They must engage in intricate reasoning, apply mathematical concepts, and tackle novel challenges. Zoller (2018) emphasised that fostering higher-order cognitive skills in students equips them with essential tools for critical thinking and lifelong learning. However, research by Nieminen, Savinainen, and Viiri (2021) found that physics assessments often fail to achieve a balanced representation of Bloom's levels, leaning heavily toward recall-based questions. This overemphasis on lower order thinking skills can hinder the development of transferable cognitive skills.

Recent findings indicate that educational results ought to be thoughtfully aligned with cognitive levels. Stelzer et al. (2020) emphasise that advancing in the study of physics requires a balanced focus on both conceptual understanding and problem-solving skills. It is asserted that the integration of both computational and conceptual enquiries enhances the inclusivity of the learning environment. Ozdemir and Isiksal (2022) examined the impact of employing various levels of Bloom's taxonomy in assessments on enhancing student engagement in the learning process. It was observed that there is an increase in motivation and cognitive effort.

In Malaysia, the implementation of OBE has highlighted the significance of Bloom's Taxonomy in the development of assessment frameworks. The Malaysian Qualifications Framework (MQF 2.0) states that Course Learning Outcomes (CLOs) should be associated with Bloom's cognitive domains (Malaysian Qualifications Agency, 2020). However, there is a lack of substantial empirical evidence regarding student performance in these domains, particularly in physics and STEM courses. In the absence of such data, educators lack the evidence necessary to enhance their test design and instructional methods.

Shanmugam and Abdullah (2021) observed that students tend to perform better on assessments when engaged in higher-order tasks, provided they receive appropriate frameworks to support their efforts. Studies conducted at Malaysian universities revealed that students perform better and demonstrate a more profound comprehension when presented with both cognitive challenges and specific guidance.

This expanding collection of research indicates a global and national trend towards increased focus on higher-order thinking skills. Nonetheless, a gap remains in assessing student performance across Bloom's levels within the realm of higher education. This investigation aims to address that gap by examining student performance on CLO1 (LOTS) and CLO2 (HOTS) in a physics assessment. This approach enhances the effectiveness of assessment strategies by utilising evidence that aligns with cognitive expectations.

Objectives Of The Study

This study aims to investigate the academic performance of undergraduate Physics students on examination questions designed at different cognitive levels of Bloom's Taxonomy. Specifically, the research focuses on comparing performance between Course Learning Outcomes (CLOs) that target lower-order thinking skills (CLO1) and those that assess higher-order thinking skills (CLO2). The study also aims to explore the implications of these performance differences for instructional and assessment design in Malaysian higher education institutions.

Research Questions

1. What is the difference in student performance between CLO1 (lower-order thinking skills) and CLO2 (higher-order thinking skills) in a Physics examination?
2. Is there a statistically significant correlation between student scores in CLO1 and CLO2?
3. What do the performance trends suggest about the alignment of assessment tasks with students' cognitive strengths and instructional practices?

METHODOLOGY

This study employed a mixed-methods approach using the sequential explanatory design, a two-phase model where quantitative data collection and analysis are followed by qualitative exploration to explain or elaborate on the statistical results (Creswell & Plano Clark, 2018). The primary objective was to assess students' performance in a university-level Physics course across different cognitive levels of Bloom's Taxonomy, and to explore the pedagogical implications of their performance trends. The study adopted a **sequential explanatory approach** (Ivankova et al., 2006), where quantitative data were collected and analyzed first, followed by qualitative data to provide deeper insights into the findings. This design was selected to ensure a comprehensive understanding of student performance while maintaining methodological rigor. Fig. 1 illustrates the research framework diagram for the study.

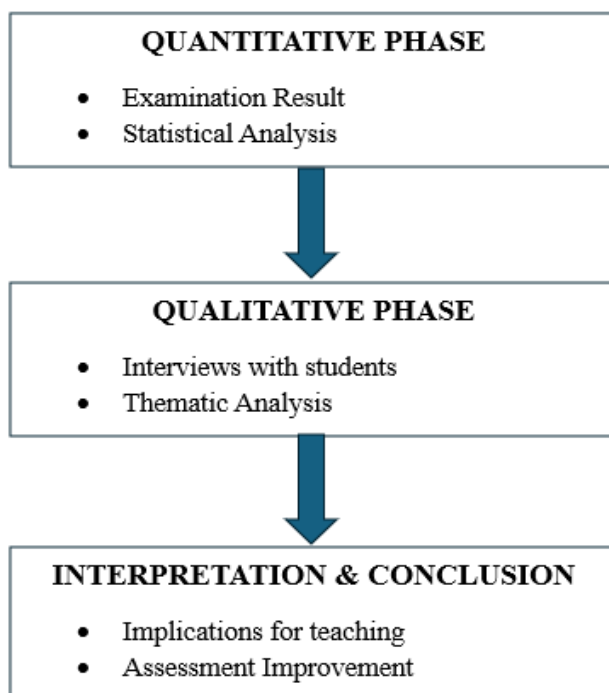


Fig. 1 Research framework diagram

In the first phase (quantitative), student scores were analysed based on their performance in examination questions aligned with two Course Learning Outcomes (CLOs):

CLO1: Targeting lower-order cognitive skills (*Remember, Understand, Apply*)

CLO2: Targeting higher-order cognitive skills (*Analyse, Evaluate, Create*)

Each section was equally weighted at 30 marks. This allowed for a direct comparison of student competency across Bloom's levels within the same cohort.

In the second phase (qualitative), semi-structured interviews were conducted with selected students and instructors to gain deeper insights into the cognitive challenges and learning strategies associated with different types of exam questions. However, this paper primarily presents the results of the quantitative analysis.

Participants

A total of **59 undergraduate students** enrolled in a core Physics course at a Malaysian higher education institution participated in this study. All students completed the same final examination paper, which included items categorized under CLO1 and CLO2. Student identities were anonymized, and data were analyzed in aggregate form to ensure confidentiality and ethical compliance.

Instrument & Assessment Design

The assessment instrument was a summative end-of-semester examination consisting of structured and problem-solving questions. The paper was divided into two main sections:

Section A (CLO1): Comprised questions testing lower order thinking skills such as definitions, basic conceptual understanding, and routine calculations. Questions focused on recall, understanding, and basic application.

Section B (CLO2): Comprised questions designed to test students' ability to analyse complex problems, evaluate scenarios, and construct solutions. Questions emphasized analysis, evaluation, and synthesis.

Each CLO carried a maximum of 30 marks. Questions were vetted by subject matter experts to ensure cognitive alignment.

Table 2 presents sample questions categorized by CLO, Bloom's level, and cognitive domain used in this study.

TABLE 2 SAMPLE EXAMINATION QUESTIONS CATEGORIZED BY CLO, BLOOMS' LEVEL AND COGNITIVE DOMAIN

CLO / Bloom's Level / Cognitive Domain	Sample Question	Mark
CLO1 Remember Lower-Order	Define nuclear fission and fusion using one example to illustrate each process.	4
CLO1 Understand Lower-Order	Explain the main differences between nuclear fission and fusion in terms of energy release and nuclear reaction conditions.	6
CLO1	A nuclear reactor produces 200 MeV of energy per	10

Apply Lower-Order	fission event. Calculate the total energy released when 0.5 mol of uranium-235 undergoes fission. Given Avogadro's number = $6.022 \times 10^{23} \text{ mol}^{-1}$	
CLO2 Analyse Higher-Order	Analyse the challenges involved in sustaining a controlled nuclear fusion reaction in a laboratory setting.	6
CLO2 Evaluate Higher-Order	Evaluate the potential of nuclear fusion reactors as replacements for fossil-fuel power plants, considering economic, environmental, and technological factors	6
CLO2 Create Higher-Order	Design a conceptual hybrid power system that integrates nuclear fission and fusion technologies, supported by a justification.	8
TOTAL MARKS		40

Data Collection & Analysis

The study used numerical data from the final exam scores of 59 undergraduate physics students. The assessment was meticulously designed to assess two distinct categories of Course Learning Outcomes. The first category of tests assessed lower order thinking skills, such as remembering, understanding, and applying basic physics concepts. The second category evaluated higher-order thinking skills, including analysing, evaluating, and creating solutions to physics-related problems. Each section of the assessment carried 40 points, contributing equally to the overall score.

Two different individuals evaluated the assessments to ensure the consistency of the scores. The evaluations conducted were notably consistent, evidenced by an inter-rater reliability coefficient (κ) of 0.85. The assessors reached a consensus regarding the validity of the scoring process. Three complementary statistical techniques were employed in quantitative analysis to rigorously investigate the data. Initially, means and standard deviations for both CLO1 and CLO2 scores were calculated as part of the descriptive analysis. To achieve a deeper insight into score distributions, the analysis additionally computed percentile distributions (25th, 50th, and 75th percentiles). To comprehensively assess student performance, both the lowest and highest scores were identified. For comparative analysis, a paired samples t-test determined whether the observed differences between CLO1 and CLO2 mean scores were statistically significant. To understand the practical significance of these differences, the study calculated effect size using Cohen's d. This provided insight into the magnitude of performance differences between the two cognitive levels. The analysis also included correlational examination using Pearson's correlation coefficient (r) to investigate the relationship between performance on lower-order and higher-order thinking skills items. The strength and direction of this relationship were interpreted using established benchmarks in educational research.

Following the quantitative analysis, the study collected qualitative data through semi-structured interviews. Researchers selected 10 students purposefully to represent high, medium, and low performers, ensuring diverse perspectives. Interviews were conducted with two physics instructors who instructed the course. This qualitative investigation examined several critical aspects of student performance. The questions were

discussed in terms of the difficulty they perceived at various levels of thought. The interviews additionally examined the methods students employed to address various types of test questions. The discussions also examined instructional elements that could influence student performance. The qualitative data was meticulously examined using the thematic analysis method outlined by Braun and Clarke (2006). The initial phase involved documenting the material and familiarising oneself with it. The audio recordings were meticulously transcribed word for word, and the transcripts were subsequently reviewed on multiple occasions to identify emerging patterns. During the coding phase, important patterns in the data were identified through the application of open coding by the researchers. The codes were subsequently categorised into groups according to potential themes that revealed broader patterns in the responses. The final phase involved revisiting these themes to ensure they precisely represented the dataset. The final themes and names were meticulously selected by the researchers to encapsulate the key insights gained from the interviews. The investigation employed quantitative and qualitative data to provide a comprehensive overview of student performance. Examining statistical patterns from the perspectives of both students and teachers enhanced their clarity. The findings indicated potential variations in test performance relative to participants' perceptions of question difficulty. This mixed-methods approach was crucial in that it enabled us to comprehend the various factors that influence performance at various levels of Bloom's taxonomy.

Multiple methods were employed to verify the results and ensure their reliability. To ensure accuracy, the quantitative results underwent extensive testing of statistical assumptions. For the qualitative components, researchers employed member checking, where participants reviewed and confirmed the accuracy of interpreted data. The study achieved methodological triangulation by converging evidence from multiple data sources, strengthening the overall reliability of conclusions. This comprehensive mixed-methods approach yielded both robust statistical evidence and rich contextual understanding of students' performance across different cognitive levels in physics assessments.

Ethical Consideration

Ethical clearance was obtained from the institutional ethics board. Informed consent was collected, and all identifying information was removed prior to analysis.

RESULTS & DISCUSSION

Table 1 shows the results of the descriptive analysis which revealed that students scored consistently higher on CLO2 ($M = 27.78$, $SD = 4.52$) compared to CLO1 ($M = 24.46$, $SD = 5.07$). The mean scores suggest that students had greater success in responding to higher-order cognitive tasks such as analysis and evaluation. In contrast, CLO1 exhibited wider variability, with student scores ranging from as low as 10 to a maximum of 41. This suggests that students had more difficulty with lower-order questions, which may have involved recall or basic comprehension. The standard deviation was also higher in CLO1, indicating greater dispersion and inconsistency in foundational knowledge among the cohort.

TABLE I DESCRIPTIVE STATISTICS FOR CLO1 AND CLO2 SCORES

Statistic	CLO1 (Lower order)	CLO2 (Higher order)
N	59	59
Mean	24.46	27.78
Standard Deviation	5.07	4.52
Minimum	10	10
25 th Percentile	21.5	26.5
Median	24.0	30.0
75 th Percentile	28.0	30.0

These findings point to a potential instructional gap in reinforcing basic concepts or a misalignment in how lower-order questions were constructed. Conversely, the narrow score range and high median in CLO2 indicate a stronger grasp of analytical and evaluative skills, potentially facilitated by instructional design or student familiarity with applied problem-solving.

To further explore the significance of these differences, a paired samples t-test was conducted where:

$$t(58) = -7.49, p < .001$$

This result confirms a statistically significant difference in students' performance between CLO1 and CLO2. The negative t-value reflects that the mean for CLO1 was significantly lower than CLO2, reinforcing the observation that students performed better in higher-order cognitive tasks.

A Pearson correlation analysis was also conducted to determine the relationship between scores in CLO1 and CLO2:

$$r = .75, p < .001$$

This strong positive correlation suggests that students who performed well in CLO1 tended to also perform well in CLO2. This implies a general consistency in students' academic capabilities across cognitive levels, although their relative performance favored the higher-order domain. must be justified, i.e. both left-justified and right-justified.

The boxplot in Fig. 2 represents the distribution of student scores for CLO1, the lower-order thinking skills and CLO2, the higher-order thinking skills for further interpretation beyond mean scores. The interquartile range (IQR) for CLO1 is notably wider compared to that of CLO2, indicating greater variability in student performance on questions targeting lower-order cognitive skills such as recall, understanding, and basic application. This spread suggests inconsistencies in students' foundational knowledge or uneven preparation for factual and routine tasks. In contrast, the CLO2 boxplot is more compact, with a narrower IQR and most student scores clustering near the upper quartile. The median score for CLO2 (30.0) is visibly higher than that of CLO1 (24.0), reinforcing the earlier statistical finding that students performed better on higher-order tasks involving analysis, evaluation, and creation. The compact shape of the CLO2 box also indicates more consistent performance among students across this section.

Additionally, CLO1 exhibited a higher number of outliers, both below and above the whiskers, suggesting that some students struggled significantly with or excelled beyond expectations on the lower-order tasks. The presence of outliers could point to differential levels of prior knowledge, question interpretation issues, or gaps in conceptual clarity. These patterns may reflect instructional alignment or misalignment. The stronger and more consistent performance on CLO2 questions suggests that classroom instruction may have been more closely aligned with analytical and evaluative thinking, potentially through problem-solving sessions, case-based discussions, or inquiry-based activities. Meanwhile, the broader performance spread and lower median in CLO1 indicate that foundational concepts may not have been emphasized or revisited with sufficient clarity, despite their assumed simplicity. This challenges the common notion that lower-order tasks are inherently easier and calls attention to the need for intentional instructional support across all cognitive domains.

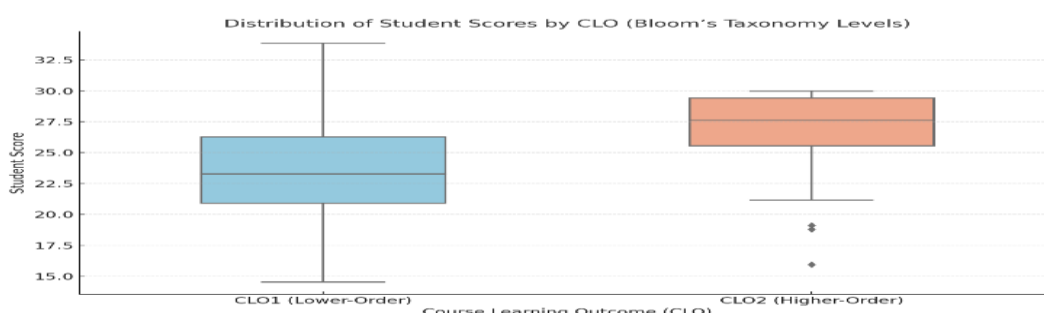


Fig. 2 Boxplot on student performance distribution for question under CLO1 and CLO2

The results of this study challenge the long-standing pedagogical assumption that lower-order cognitive skills, such as recalling facts or basic understanding, are easier and thus produce better student outcomes. Contrary to expectation, students demonstrated significantly higher performance on CLO2 tasks, which assessed higher-order thinking skills like analysis, evaluation, and synthesis. This finding highlights a potential mismatch between the perceived cognitive difficulty of assessment items and how students engage with them. One possible explanation is the evolving instructional paradigm in Physics education. Increasing use of problem-based and inquiry-driven learning approaches in Malaysian higher education may have better prepared students to tackle analytical and evaluative tasks. These approaches frequently encourage analytical reasoning, teamwork in addressing challenges, and the practical use of concepts in real-life situations and abilities that are more in line with CLO2.

The motivation and engagement of students could also be influential factors. Individuals might perceive tasks demanding advanced cognitive skills as more intellectually engaging and pertinent, potentially resulting in greater mental involvement. Conversely, activities centred around CLO1 may have been perceived as overly simplistic or disconnected from practical problem-solving, potentially leading to diminished interest and effort from participants. It is essential to consider the configuration of the test. Although the CLO1 questions were designed to assess fundamental comprehension, they might have appeared more challenging than intended due to complex phrasing, abstract contexts, or a lack of alignment with the instructional material presented in class. Conversely, CLO2 questions may have been more comprehensible or resembled previously addressed problems.

The strong positive correlation ($r = .75$) between CLO1 and CLO2 scores suggests that students who performed well in one cognitive domain were likely to do well in the other. This reinforces the notion that general academic capability and mastery of content play a central role across all levels of cognitive assessment. However, the significant performance gap highlights an opportunity to refine instructional and assessment strategies to ensure foundational skills are as well-supported as higher-order capabilities. These findings are consistent with those of Nieminen et al. (2021) and Zoller (2018), who reported that students often rise to the challenge of higher-order questions when they are properly scaffolded. Thus, educators should be encouraged to integrate such questions more confidently into assessments while also ensuring robust support for lower-order skills.

Ultimately, the results point to the need for a balanced approach in curriculum design—one that equally fosters foundational knowledge and advanced cognitive processing. Comprehensive cognitive development across Bloom's Taxonomy is critical not only for academic success but also for preparing graduates with the analytical skills required in a knowledge-driven economy.

Limitation

While the findings of this study offer valuable insights into student performance across Bloom's cognitive levels, several limitations must be acknowledged. First, the study was conducted within a single Malaysian higher education institution and involved only one cohort of physics students. This context-specific focus may limit the generalizability of the results to other institutions or disciplines. Additionally, the nature of the assessment tasks, particularly those aligned with CLO2 may have been shaped by the institutional assessment culture or instructional practices, introducing potential bias. Although the classification of exam questions into Bloom's Taxonomy levels was reviewed by subject matter experts, the process inherently involves a degree of subjectivity, which could result in classification bias. Furthermore, while the sample size of 59 students is adequate for statistical analysis, it may not fully represent the broader population of undergraduate physics students in Malaysia. These limitations should be considered when interpreting the study's conclusions and planning future research.

CONCLUSION

This study suggests that undergraduate Physics students at a Malaysian university perform significantly better on assessments designed to evaluate higher-order thinking skills, such as analysis and evaluation, compared to those focused on lower-order skills, including remember and understanding. The findings indicate that students

might exhibit a greater interest in or enhanced readiness for complex cognitive tasks. This is likely due to instructional approaches that emphasise problem-solving, practical application, and analytical reasoning. The variation in performance suggests that the notion of lower-order tasks being consistently easier or more accessible is less plausible. This indicates a necessity to reevaluate our approaches to teaching and assessing fundamental skills, ensuring that every student receives equitable support across all cognitive domains.

RECOMMENDATION

Based on the findings, several strategic recommendations are made to improve evaluation methods and promote cognitive skill development in Physics education. It is crucial to encourage educators to include a variety of cognitive tasks from basic to advance in their assessments. Students are assessed on their critical thinking, judgement, and information synthesis skills as well as their ability to recall facts and apply basic concepts. A comprehensive question-writing method would reveal a student's cognitive abilities. Cognitive objectives should match lesson methods. Active learning methods like inquiry-based learning, case-based instruction, and conceptual modelling can improve basic and advanced cognitive skills. Additionally, students must be clearly informed of assessment expectations. Guidelines, examples, and detailed instructions can help students overcome cognitive challenges and become more aware of their thought processes. Test questions must be reviewed and validated regularly to match course outcomes and cognitive levels. Teachers and curriculum committees must regularly review test questions for clarity, appropriateness, and fairness. This continuous evaluation process helps identify questions that are overly complicated due to unclear wording or framing. Faculty enhancement programs should support educators. Bloom's Taxonomy, effective item construction, and data-informed instruction help educators create better assessments and understand student performance better. Including student feedback in the evaluation process can also reveal students' perceptions of question difficulty, relevance, and clarity. Organised feedback forms or reflections after an exam can influence assessment material changes. Future studies should be more qualitative and extensive. Interviews, class observations, and reflective journal analysis can reveal students' thought processes when answering questions. Improved comprehension may enable focused strategies that balance Bloom's cognitive hierarchy. These strategies help educational institutions match instructional and assessment methods to students' cognitive growth needs. This will improve academic performance in all fields.

ACKNOWLEDGMENT

The authors wish to express their appreciation to the Physics students who took part in this study. Their readiness to provide insights into their academic performance and engage in educational inquiry has significantly enhanced our comprehension of cognitive skill development within the realm of higher education.

REFERENCES

1. Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.
2. Biggs, J., & Tang, C. (2011). Teaching for quality learning at university (4th ed.). McGraw-Hill.
3. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
4. Caballero, M. D., & Wilcox, B. R. (2017). Analyzing the role of conceptual scaffolding in solving physics problems. *Physical Review Physics Education Research*, 13(1), 010121. <https://doi.org/10.1103/PhysRevPhysEducRes.13.010121>
5. Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.
6. Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64-74. <https://doi.org/10.1119/1.18809>
7. Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods*, 18(1), 3–20.
8. Malaysian Qualifications Agency. (2020). Malaysian Qualifications Framework 2.0 (MQF 2.0).

<https://www2.mqa.gov.my/mqf2>

9. Nieminen, P., Savinainen, A., & Viiri, J. (2021). Assessing the quality of physics questions with Bloom's taxonomy. *European Journal of Physics*, 42(3), 035703. <https://doi.org/10.1088/1361-6404/abef19>
10. Ozdemir, A. S., & Isiksal, M. (2022). The effect of integrating Bloom's taxonomy into assessment: A focus on student engagement. *International Journal of Science and Mathematics Education*, 20(4), 765–784.
11. Redish, E. F. (2003). *Teaching physics with the physics suite*. Wiley.
12. Roshayanti, Y., Putri, A. E., & Nurhikmah, H. (2022). Enhancing higher-order thinking skills through structured physics assessments: A case study in Indonesian senior high schools. *Journal of Science Education Research*, 6(1), 22–35.
13. Shanmugam, K., & Abdullah, M. A. (2021). Cognitive scaffolding and HOTS development in Malaysian universities. *Malaysian Journal of Learning and Instruction*, 18(2), 45–62.
14. Stelzer, T., Gladding, G., Mestre, J. P., & Brookes, D. T. (2020). Comparing the efficacy of conceptual and computational physics problems. *Physical Review Physics Education Research*, 16(2), 020112. <https://doi.org/10.1103/PhysRevPhysEducRes.16.020112>
15. Zoller, U. (2018). Higher-order cognitive skills (HOCS) in science education: Current trends and future directions. *Science Education International*, 29(1), 3–12.