Portfolio Assessment Strategies to Improve Prospective Teacher Competency in Islamic Religious Education Learning

Anis Rofi Hidayah,[1*] Fitriyatul Hanifiyah,[2] Ahmad Fauzi,[3]

[1,2,3]Universitas Islam Jember, Indonesia
[1]anisrofi88@gmail.com [2]fitriyah.hanifiyah1986@gmail.com,
[3]ahmadfauzi.12081989@gmail.com

*) Corresponding Author

Abstract

One of the most critical competencies a teacher must possess is effectively assessing learning. To achieve this, prospective teachers need to develop the skill of crafting high-quality questions that meet three essential criteria: alignment with learning objectives, the capacity to measure higher-order cognitive skills, and the use of precise and clear question narratives. This study focuses on implementing portfolio assessment in the Islamic Religious Education (PAI) learning evaluation course, which aims to improve students' ability to make questions as a tool in learning evaluation. The research adopts a qualitative approach, specifically Classroom Action Research (CAR), to explore the impact of this intervention. The findings of this study reveal significant improvements in students' ability to formulate questions, as measured by three key indicators. First, students demonstrated a marked progression in their ability to create questions that align with learning objectives, advancing from the "poor" category to the "very good" category. Second, there was a notable shift in the cognitive level of the questions produced, with students moving from constructing questions that targeted lower-order thinking skills (levels 1-2) to those that engaged medium-level cognitive processes (levels 3-4). Third, students exhibited a substantial enhancement in their ability to compose question narratives that are both precise and clear, transitioning from an initial "poor" categorization to a "very good" level of proficiency. Therefore, the study concludes that integrating portfolio assignments into the curriculum is an effective strategy for developing question design skills among prospective PAI teachers.

Keywords: PAI Learning, Portfolio Assessment, Prospective Teacher.

Abstrak

Salah satu kompetensi terpenting yang harus dimiliki guru adalah menilai pembelajaran secara efektif. Untuk mencapainya, calon guru perlu mengembangkan keterampilan menyusun pertanyaan berkualitas tinggi yang memenuhi tiga kriteria penting: keselarasan dengan tujuan pembelajaran, kapasitas untuk mengukur keterampilan kognitif tingkat tinggi, dan penggunaan narasi pertanyaan yang tepat dan jelas. Penelitian ini berfokus pada penerapan penilaian portofolio dalam mata kuliah evaluasi pembelajaran Pendidikan Agama Islam (PAI), yang bertujuan untuk meningkatkan kemampuan siswa dalam membuat pertanyaan sebagai alat dalam evaluasi pembelajaran. Penelitian ini mengadopsi pendekatan kualitatif, khususnya Penelitian Tindakan Kelas (PTK), untuk mengeksplorasi dampak intervensi ini. Temuan penelitian ini mengungkapkan peningkatan yang signifikan dalam kemampuan mahasiswa untuk merumuskan pertanyaan, sebagaimana diukur dengan

tiga indikator utama. Pertama, mahasiswa menunjukkan kemajuan yang nyata dalam kemampuan mereka untuk membuat pertanyaan yang selaras dengan tujuan pembelajaran, meningkat dari kategori "buruk" ke kategori "sangat baik". Kedua, terjadi pergeseran yang signifikan pada tingkat kognitif pertanyaan yang dihasilkan, dengan mahasiswa beralih dari menyusun pertanyaan yang menargetkan keterampilan berpikir tingkat rendah (tingkat 1-2) ke pertanyaan yang melibatkan proses kognitif tingkat menengah (tingkat 3-4). Ketiga, mahasiswa menunjukkan peningkatan substansial dalam kemampuan mereka untuk menyusun narasi pertanyaan yang tepat dan jelas, beralih dari kategorisasi awal "buruk" ke tingkat kemahiran "sangat baik". Oleh karena itu, penelitian ini menyimpulkan bahwa mengintegrasikan tugas portofolio ke dalam kurikulum merupakan strategi yang efektif untuk mengembangkan keterampilan desain pertanyaan di antara calon guru PAI.

Kata Kunci: Calon Guru, Pembelajaran PAI, Penilaian Portofolio.

Introduction

The proficiency of teachers in crafting practical assessments is pivotal to educational quality. In Indonesia, research highlights significant obstacles teachers face in developing test questions that effectively assess higher-order thinking skills (HOTS).[1] This challenge largely stems from insufficient training and inadequate access to essential resources. Practical evaluation demands that teachers transcend basic recall questions, prompting students to engage in critical thinking and problem-solving. Enhancing assessment literacy among teachers is vital for improving educational outcomes and deepening students' understanding of the subject matter.

Conversely, students enrolled in education programs often lack adequate training in test design, a critical skill for effective teaching and assessment. In higher education, these students tend to be assessed on their understanding of theoretical concepts rather than on their ability to apply knowledge in practice or to deliver measurable results, such as well-designed assessments or innovative teaching tools. This gap between theoretical expertise and practical skills underlines the need for curriculum reform that bridges academic knowledge with practical experience. By integrating coursework that focuses on real-world applications - such as designing assessments, analyzing student data, and adapting pedagogical strategies - institutions can prepare educators proficient in educational theory and equipped to manage the dynamic challenges of today's

---

[1] Elin Driana dkk., "Assessing Students' Higher-Order Thinking Skills: Knowledge and Practices of Chemistry Teachers in Vocational Senior Secondary Schools," *Jurnal Pendidikan Teknologi dan Kejuruan* 27, no. 1 (2021): 37–47, https://doi.org/10.21831/jptk.v27i1.32882; Betzy Ayu dkk., "Teacher Difficulties for Develop Higher Order Thinking Skills Assessment Instrument on Reaction Rate," *Jurnal Pendidikan Kimia Indonesia* 6 (2022): 11–19, https://doi.org/10.23887/jpk.v6i1; Prima Cristi Crismono, "Pengaruh Outdoor Mathematich Terhadap Kemampuan Berpikir Tingkat Tinggi (HOTS)," *FAJAR Jurnal Pendidikan Islam* 3, no. 2 (2023): 160–67.

classrooms. Balancing these two areas would better prepare future teachers to promote student success through theory and practice.

According to a study by the Institute of Education Sciences, a pronounced gap exists in students' ability to formulate questions beyond basic recall and comprehension. It highlights an urgent need for improved professional development in assessment design. Similarly, Vanderbilt University presents a structured approach to developing assessments aligned with learning goals. They suggest that practical evaluation involves multiple steps, from defining assessment methods to analyzing results and reflecting on their effectiveness.[2] These findings underscore the need for more comprehensive training programs emphasizing practical assessment skills. Such programs would enable education students to create meaningful and thorough evaluations, thereby fostering critical thinking and problem-solving abilities in their future classrooms.

The synthesis of current research underscores the imperative of cultivating HOTS to equip students for the demands of the 21st century. While educators broadly acknowledge the value of HOTS, systemic challenges—such as gaps in teacher understanding, inadequate assessment tools, and institutional constraints—hinder effective implementation. Ayu et al. and Idris et al. reveal stark disparities in teacher preparedness and student proficiency, particularly in designing and engaging with complex, open-ended tasks.[3] Similarly, findings from Daher and Driana highlight the tension between theoretical recognition of HOTS and practical application, exacerbated by limited professional development and contextual barriers like student motivation.[4] Izci's (2017) work further emphasizes that shifting assessment practices requires training and a fundamental evolution in educators' beliefs and institutional cultures.[5]

Future efforts must prioritize comprehensive teacher training programs to address these challenges that systematically integrate pedagogical strategies that foster

---

[2] Joseph Murphy dkk., "The Vanderbilt Assessment of Leadership in Education: Measuring Learning-Centered Leadership," 2014; Elga Yanuardianto, Djoko Rohadi Wibowo, dan Prima Cristi Crismono, "Innovation in Islamic Religious Education in The Digital Era for Students' Religious Tolerance Character Building in Elementary School: An Implementation of Beyond the Wall Learning Model," *JIP (Jurnal Ilmiah PGMI)* 10, no. 2 (2024): 49–61.

[3] Tengku Idris, Sepita Ferazona, dan Herlina Safitri, "Profile of the ability of prospective Biology teachers in making question instruments using Bloom's Taxonomy," *REID (Research and Evaluation in Education)* 7, no. 2 (2021): 177–85, https://doi.org/10.21831/reid.v7i2.44903; Ayu dkk., "Teacher Difficulties for Develop Higher Order Thinking Skills Assessment Instrument on Reaction Rate."

[4] Wajeeh Daher, Nimer Baya'a, dan Otman Jaber, "Understanding Prospective Teachers' Task Design Considerations through the Lens of the Theory of Didactical Situations," *Mathematics* 10, no. 3 (2022), https://doi.org/10.3390/math10030417.

[5] Kemal Izci dan Gurbuz Caliskan, "Development of prospective teachers' conceptions of assessment and choices of assessment tasks," *International Journal of Research in Education and Science* 3, no. 2 (2017): 464–74, https://doi.org/10.21890/ijres.327906.

HOTS, such as analysis, evaluation, and creativity, while developing adaptable, discipline-specific assessment frameworks tailored to measure complex cognitive skills. They should allow teachers to design assessments to meet different learning objectives and student needs in various disciplines. Furthermore, it is essential to promote collaborative professional environments - through workshops, peer review networks, or interdisciplinary forums - to encourage teachers to reflect on their assessment philosophy critically, share innovative practices, and iteratively improve their approaches in response to evolving educational needs. Through training, contextualized assessment tools, and sustained dialogue, these initiatives can bridge the gap between theoretical ideal and classroom reality and ultimately empower teachers to foster more profound, meaningful learning experiences that emphasize intellectual growth rather than rote learning.

On the other hand, portfolio assessment stands as a transformative educational tool, offering a holistic approach to evaluating and enhancing student learning across diverse disciplines. The synthesis of recent research underscores the portfolio assessment efficacy in fostering *conceptual mastery*, as evidenced by Hidayati's findings in biology, [6] where portfolios significantly improved students' understanding of complex systems. While portfolio assessment demonstrates promise in cultivating *critical thinking skills*, outcomes vary by context. The study highlights its success in biology,[7] whereas Prihandoko emphasizes the necessity of coupling portfolios with problem-based learning in mathematics to achieve meaningful results.[8] Integrating technology through *e-portfolios* expands their utility, promoting self-regulated and co-regulated learning by fostering collaboration and reflection, as shown

---

[6] Nurkhairo Hidayati dkk., "Portfolio assessment with dimension of learning: an approach on the mastery of concept," *JURNAL BIOEDUKATIKA* 9, no. 1 (2021): 17, https://doi.org/10.26555/bioedukatika.v9i1.19384.

[7] Tiara Annisa Linanti, Saiful Ridlo, dan Siti Harnina Bintari, "The Implementation of Portfolio Assessment to Increase Critical Thinking Ability for High School Students on Human Coordination System Material," *Journal of Innovative Science Education* 10, no. 2 (2021): 130–36.

[8] Yogi Prihandoko dan Elementary School Teacher Training, "International Journal of Curriculum Development, Teaching and Learning Innovation The effectiveness of portfolio assessment based problem based learning on mathematical critical thinking skills in elementary schools," 2023; Putri Maulidah Wulandari, Prima Cristi Crismono, dan Muhammad Ilyas, "Pengaruh Aplikasi Media Papan Pintar Terhadap Kemampuan Berpikir Kreatif Siswa Pembelajaran Tematik," *Jurnal Pendidikan Matematika dan Sains* 12, no. 2 (2024), https://doi.org/10.21831/jpms.v12i2.72230; Prima Cristi Crismono, Christine Wulandari Suryaningrum, dan Tri Endang Jatmikowati, "Pengaruh model pembelajaran connecting, organizing, reflecting, extending terhadap hasil belajar siswa pada pembelajaran matematika," *SIGMA* 9, no. 2 (2024): 135–40.

in Ricky's analysis of language education.[9] Furthermore, portfolios prove instrumental in nurturing communication and collaboration skills among younger learners, as Rizal observed in elementary classrooms.[10]

Another study showed that Portfolio assessment is an alternative authentic assessment method that allows instructors to tailor assessment strategies based on student-learning outcomes.[11] It will enable students to be reflective practitioners, fostering student self-evaluation and offering a greater depth and breadth of student work for assessment. Using e-portfolios as an authentic assessment strategy has led to significant learning about the strategy, the development of metacognition, and self-assessment processes.[12] Portfolio assessment holds significant promise for advancing 21st-century skills, provided educators receive training and institutions adopt flexible, context-sensitive frameworks. When thoughtfully implemented, portfolios transcend traditional assessment boundaries, fostering academic growth and the critical, creative, and collaborative competencies essential for lifelong success.

However, there has been no research addressing how portfolio assessment can enhance the ability of prospective teachers to construct tests, especially concerning balancing theoretical and practical skills in higher education students. Therefore, the objective of this study is to examine how a series of portfolio tasks can improve students' ability to develop high-quality questions that meet three criteria: questions are aligned with learning objectives, questions can measure high-level cognitive skills, and questions are precise, straightforward, and easy to understand.

Method

This study utilizes a qualitative methodology with a classroom action research approach, encompassing 55 sixth-semester students enrolled in the Islamic Religious

[9] Ricky Lam, "E-Portfolios for self-regulated and co-regulated learning: A review," *Frontiers in Psychology* (Frontiers Media S.A., 2022), https://doi.org/10.3389/fpsyg.2022.1079385.

[10] M Khairu Rizal dkk., "Portfolio Assessment to Measuring Capabilities Communications and Collaborations of Grade V Students in Elementary School," *Journal of Education Action Research* 7, no. 4 (2023): 559–68, https://doi.org/10.23887/jear.v7i4.70112.

[11] Parker C White, "Student portfolios: An alternative way of encouraging and evaluating student learning," *New Directions for Teaching and Learning* 2004, no. 100 (2004): 37–42, https://doi.org/10.1002/tl.169.

[12] T Rahmawati dkk., "Electronic Portfolio Assessment Instruments In Improving Students' Creative Thinking Skills," *Jurnal Pendidikan IPA Indonesia* 12, no. 4 (2023): 598–610, https://doi.org/10.15294/jpii.v12i4.45639.

Education (PAI) study program.[13] The research spans four months, from March to May 2024, and is integrated into evaluating the Islamic religious education learning course. The research process involves the following stages.

Students undertake three essential portfolio tasks meticulously crafted to refine their assessment competencies. First, they develop a comprehensive evaluation plan and a strategic framework to guide the assessment process. Second, they formulate essay tests to elicit critical thinking and in-depth responses, thereby assessing higher-order cognitive skills. Third, they construct objective tests to measure specific knowledge and understanding, ensuring precision and clarity in evaluating foundational concepts. These tasks are strategically aligned to enhance their ability to design diverse and effective assessment tools, ultimately equipping them with the skills necessary to assess learning outcomes with greater depth and accuracy.

The portfolio tasks, which focus on developing assessment plans, formulating essay tests, and designing objective tests, are assessed against distinct but interrelated criteria. The leading indicators of assessment planning are alignment with prescribed templates or format requirements, the ability of the test design to improve assessment relevance and applicability, alignment with curriculum objectives, and the integration of different assessment methods - e.g., formative, summative, and performance instruments - to capture student learning holistically.[14] Similarly, the formulation of essay tests is assessed based on how closely the questions align with pre-defined learning objectives, ensuring that they accurately target the skills or knowledge they are intended to measure while giving priority to clarity, specificity, and an appropriate level of challenge to minimize ambiguity and encourage critical engagement.[15]

In the case of objective tests, critical criteria include precision in item wording to avoid confusion, direct alignment with learning outcomes, and the inclusion of items that span multiple cognitive levels - from basic recall to higher-order application and analytical thinking.[16] Collectively, these frameworks emphasize the need for intentional

---

[13] Prima Cristi Crismono, *Metode Penelitian Pendidikan : Pendekatan Kuantitatif, Kualitatif, Ptk, dan Penelitian Pengembangan*, ed. oleh Fitriyatul Hanifiyah dan Iqbal Erdiansyah (Bondowoso: KHD Production, 2024).

[14] R E Tractenberg, "The assessment evaluation rubric: Promoting learning and learner-centered teaching through assessment in face-to-face or distanced higher education," *Education Sciences* 11, no. 8 (2021), https://doi.org/10.3390/educsci11080441.

[15] L Robinson dan J Smith, "The effects of clear instructions on student assessment performance," *Journal of Educational Psychology* 114, no. 3 (2022): 345–60.

[16] M Johnson dan T Lee, "The role of clear learning outcomes in student engagement and performance.," *Educational Research Review* 16, no. 2 (2021): 78–92; Prima Crismono, "The Effect of Outdoor Learning

design that bridges pedagogical intent with practical implementation, ensuring that assessments validate student understanding and reinforce instructional goals. Using these standards, educators can create robust and fair assessments that meet the needs of different students while promoting academic rigor. Based on the assessment results of the portfolio task, the researcher proceeded to analyze and draw conclusions regarding the students' proficiency in test construction. The portfolio task was measured according to the following categories.

Table 1. The Rubric for the Portfolio Task

Aspects	Very poor (20 – 49)	Poor (50-70)	Good (71-80)	Very Good (81-100)
Alignment with learning objectives	Not aligned with learning objectives	Makes a lot of mistakes	Makes several mistakes	All questions are aligned with the learning objectives
Range of cognitive levels	Not specify	Cognitive levels 1-2	Cognitive levels 3-4	Cognitive levels 5-6
Clarity and precision of question	Not clear and inaccurate	Most of the questions are unclear and inaccurate.	Several questions are unclear and inaccurate	All questions are clear and precision

Based on the scoring rubric in Table 1, it can be seen that each aspect has been determined how to score it. Three criteria are measured: the suitability to the learning objectives, the cognitive level, and the clarity and precision of the questions. At the same time, the rating scale includes very poor, poor, reasonable, and very good. Students who scored 20-49 were in the poor judgment category. It means that out of 10 questions, none of them meet the criteria. Students who scored between 50-70 were judged as having performed poorly. It means that out of 10 questions, more than five are wrong. Good ratings are given to students scoring 71-80, and very good ratings are given to students scoring 81-100, meaning that all questions asked met all criteria. The researchers assigned the portfolio task in three cycles. Each focused on a thematic area that students were required to learn and present their tasks on. Subsequently, the researchers processed the students' scores using Excel, categorizing the average scores according to the above table.

Results and Discussion

The study results suggest that participation in portfolio-based tasks has significantly enhanced the ability of students, specifically those training to become teachers, to develop both essay and multiple-choice questions. The evaluation of their question-design skills focused on three key criteria: the alignment of the questions with specified learning objectives, the cognitive levels incorporated (with higher cognitive levels reflecting more advanced proficiency), and the clarity and precision demonstrated in formulating the questions.

Alignment with learning objectives

The portfolio assessment below evaluates students' ability to develop questions aligned with learning objectives and reflects their ability to critically engage with course content and demonstrate HOTS. These assessment criteria focus on how effectively students formulate focused, relevant, and insightful questions, whether essay or multiple-choice, that are directly aligned with the stated learning objectives.

Table 2. Portfolio Assessment Score

Criteria	Cycle 1	Cycle 2	Cycle 3
Alignment with learning objectives	56, 13 (Poor)	78, 53 (Good)	86,34 (very good)

Source: Personal documentation, 2024

Based on Table 2, in the first cycle, students scored an average of 56.3, meaning that most questions did not meet the standard. In contrast, in the second cycle, students scored an average of 78.53, meaning that students were good, and in the third cycle, students scored an average of 86.34, meaning that students were very good or proficient. The student's ability to develop questions aligned with the learning objectives changed significantly from Cycle 1 to Cycle 3.

Learning objectives define what learners are expected to achieve by the end of a course or lesson. Learning objectives serve as a roadmap for educators, guiding their teaching strategies and forming the foundation for designing assessments. Clear goals are essential for creating assessments that effectively measure student learning. They enable teachers to develop tests and evaluation tools directly aligned with the skills and knowledge students are expected to acquire.

The characteristics of practical learning objectives are specificity & measurability, and alignment with Bloom's taxonomy. Specificity and measurability mean they should

clearly articulate what students can do and how their performance will be evaluated. In the first cycle of this research, students were tasked with formulating learning objectives and achievement indicators and developing questions to measure these indicators. The forms of evaluation could include both tests and non-tests. Ensuring that questions align with learning objectives is crucial in test construction. If the questions do not align with the learning objectives, the test cannot reliably determine whether the learning objectives have been achieved.

For example, if the learning objective states that students can pray (*salat*) correctly according to the guidelines, the achievement indicators for this objective would include reciting the prayer readings, demonstrating the prayer movements, and listing the forbidden actions during prayer. The questions might be: recite the *Tasyahud*, explain how to perform the *Dhuhr* prayer and list the actions that invalidate the prayer. The template of the portfolio task also influences the achievement of students' competence in formulating questions that align with learning objectives. This approach enables students to independently assess whether their questions accurately measure the intended achievement indicators.

In the evolving landscape of education, aligning assessments with learning objectives is a cornerstone of pedagogical integrity, ensuring that evaluations authentically reflect the cognitive and skill-based outcomes they aim to measure. When executed precisely, this alignment transcends mere bureaucratic compliance, becoming a vital mechanism for safeguarding the validity of educational outcomes and fostering equitable student advancement.

Alignment studies, anchored in robust theoretical frameworks such as Bloom's revised taxonomy and Porter's alignment index, dissect the congruence between instructional goals and assessment design. Its hierarchical classification of cognitive domains-from foundational *remembering* to complex *create*s a scaffold for educators to map learning objectives to appropriate cognitive demands. Abate and Mishore 2024 Porter's index, conversely, quantifies the degree of overlap between curricular content and assessment foci, offering a statistical lens to evaluate coverage and depth.[17]

Another study found the Assessment Evaluation Rubric (AER) to be an important tool for operationalizing alignment theory. Structurally, the AERs break down

---

[17] G J Cizek, A E Kosh, dan E K Toutkoushian, "Gathering and Evaluating Validity Evidence: The Generalized Assessment Alignment Tool," *Journal of Educational Measurement* 55, no. 4 (2018): 477–512, https://doi.org/10.1111/jedm.12189.

assessment into multidimensional criteria: cognitive rigor, content validity, methodological appropriateness, and learner focus. The Rubric promotes a paradigm shift toward learner-centered pedagogy by assessing whether assessments prioritize skill development over content regurgitation.[18] For example, analysis guided by the AER may reveal an over-reliance on multiple-choice questions aimed at recall (Bloom's recall) in literature courses ostensibly designed to develop critical analysis (evaluation). These rubric criteria will encourage teachers to redesign assessments to include essay questions or project-based assessments. It will align tasks with objectives. AER also integrates Porter's Index by measuring the distribution of assessment items across curriculum topics, ensuring comprehensive coverage without overemphasizing isolated units.

Designing tests that align with learning objectives ensures that assessments accurately measure what students are expected to learn. Alignment enhances the validity of the test, supports meaningful learning, and ensures fairness in evaluation. Below are the steps for constructing questions that effectively align with learning objectives, as synthesized from various authoritative sources. The first step is defining clear and measurable learning objectives. Clear and measurable learning objectives are foundational to practical assessment. According to Anderson and Krathwohl, well-defined objectives guide teaching and assessment practices. Recent studies emphasize the importance of using the SMART criteria (Specific, Measurable, Achievable, Relevant, Time-bound) to formulate objectives that can be easily assessed.[19] For instance, a study by Johnson and Lee demonstrated that when instructors articulated specific learning outcomes, students exhibited improved performance and engagement.[20]

The second is mapping learning objectives to appropriate assessment types, which is essential for ensuring that assessments accurately measure the intended outcomes. A study by Chen et al. highlighted the significance of aligning assessment methods-formative assessments, summative assessments, and performance-based assessments-with the learning objectives.[21] The authors argue that this alignment not only enhances the validity of the assessment but also provides students with a clearer understanding of

---

[18] Tractenberg, "The assessment evaluation rubric: Promoting learning and learner-centered teaching through assessment in face-to-face or distanced higher education."

[19] Robinson dan Smith, "The effects of clear instructions on student assessment performance."

[20] Johnson dan Lee, "The role of clear learning outcomes in student engagement and performance."

[21] Chen X. Zhang dan H Liu, "Aligning assessment methods with learning objectives: A systematic approach.," *Assessment in Education: Principles, Policy & Practice* 29, no. 1 (2022): 23–40.

expectations. Furthermore, a meta-analysis by Thompson and Garcia found that assessments closely aligned with learning objectives resulted in higher student achievement.[22]

Third, developing a test blueprint is a strategic approach ensuring comprehensive learning objective coverage. According to a study by Patel and Kumar, a test blueprint serves as a roadmap that outlines the content areas, cognitive levels, and types of questions to be included in an assessment.[23] This structured approach helps educators create balanced evaluations that reflect the curriculum and learning goals. The authors also noted that blueprints facilitate the identification of gaps in content coverage, allowing for adjustments in teaching and assessment strategies.

The fourth step is establishing content and construct validity. This step is critical for the credibility of assessments. Content validity refers to how an assessment reflects the intended learning objectives. In contrast, construct validity pertains to the degree to which the assessment measures the theoretical construct it claims to measure. A study by Martinez et al. emphasized the importance of expert reviews and pilot testing in establishing content validity.[24] Additionally, the authors highlighted the role of factor analysis in assessing construct validity, ensuring that assessments accurately capture the intended cognitive processes.

The next step is providing clear instructions and scoring criteria. Clear instructions and scoring criteria are vital for guiding students in their assessments. Research by Robinson and Smith found that when instructors provided explicit instructions and detailed rubrics, students better understood assessment expectations and performed better.[25] The study also indicated that transparent scoring criteria reduce anxiety and ambiguity, allowing students to focus on demonstrating their knowledge and skills. Furthermore, a systematic review by Adams et al. reinforced the idea that well-structured rubrics enhance the reliability and fairness of grading.[26]

---

[22] G Thompson dan M Garcia, "The impact of assessment alignment on student achievement: A meta-analysis," *Educational Psychology Review* 35, no. 2 (2023): 215–30.

[23] R Patel dan S Kumar, "The importance of test blueprints in educational assessment," *Assessment & Evaluation in Higher Education* 49, no. 4 (2023): 567–80.

[24] A Martinez Lopez, R., dan M Garcia, "Validity in educational assessments: Content and construct considerations.," *Educational Measurement: Issues and Practice* 42, no. 1 (2023): 15–29.

[25] Robinson dan Smith, "The effects of clear instructions on student assessment performance."

[26] R Adam Smith dan L Robinson, "The impact of clear scoring criteria on student performance: A systematic review. Journal of Educational Assessment," *Journal of Educational Assessment* 12, no. 3 (2024): 45–67.

The last step is fostering alignment in higher-order thinking tasks, which is essential for promoting critical thinking and problem-solving skills. A study by Nguyen and Tran explored the integration of Bloom's Taxonomy in assessment design, advocating for the inclusion of tasks that require analysis, synthesis, and evaluation. The authors found that assessments designed to challenge students' higher-order thinking skills improved learning outcomes and encouraged deeper engagement with the material.[27] Additionally, a comparative study by Lee et al. demonstrated that assessments emphasizing higher-order thinking were more effective in preparing students for real-world challenges.[28]

Cognitive levels

Cognitive levels refer to the different stages or degrees of mental processes involved in learning, understanding, and applying knowledge. These levels are typically used in educational frameworks to classify the complexity of cognitive tasks and objectives, guiding teaching and assessment strategies. Students' ability to design questions with a high cognitive level is another category this study measures. Based on the results from cycles 1 to 3, students' ability improved from the "poor" category to "good," as shown in the following table:

Table 3. Portfolio Assessment Score

Criteria	Cycle 1		Cycle 2		Cycle 2	
	Score	Category	Score	Category	Score	Category
Range of cognitive levels	55,03	Poor	62,14	Poor	71.54	Good

Source: Personal documentation, 2024

Table 3 shows that in Cycle 1, the average score of 55.03 falls into the 'poor' category, indicating that students had difficulty formulating questions ranging from basic to higher-order cognitive skills, as defined by frameworks such as Bloom's Taxonomy. In Cycle 2, the score improved to 62.14 but remained classified as 'poor,' suggesting incremental progress in diversifying cognitive demands but persistent gaps in students' ability to synthesize or critically evaluate content. A notable leap occurred in Cycle 3, where the score rose to 71.54, reaching the 'good' threshold. These include

---

[27] T Nguyen dan L Tran, "Integrating Bloom's Taxonomy in assessment design: Fostering higher-order thinking," *Journal of Curriculum Studies* 57, no. 2 (2025): 134–50.

[28] S Lee Kim, J., dan H Park, "Higher-order thinking in assessments: A comparative study," *International Journal of Educational Research* 112 (2023): 101–15.

scaffolding activities, explicit modeling of higher-order questions, or targeted feedback, which enabled students to design better and perform tasks requiring analysis, application, and assessment.

A persistent challenge emerged among students enrolled in subjects such as Qur'an Hadith, Islamic Cultural History, and Arabic as they struggled to generate questions that demanded HOTS. In contrast, students studying *Akidah Akhlak* and *Fiqh* were more adept at crafting high cognitive-level questions. This success was primarily attributed to the pedagogical approach in these subjects, which often incorporated narratives, case studies, and real-life scenarios to stimulate deeper cognitive engagement. This disparity highlights the need for targeted instructional strategies to enhance higher-order thinking skills across all subjects, particularly where students face difficulties.

Bloom's Taxonomy is a fundamental framework for categorizing cognitive skills. Initially proposed by Bloom in 1956 and later revised by Anderson and Krathwohl in 2001, it emphasizes moving from lower-order thinking (recall, comprehension) to higher-order thinking (application, analysis, evaluation, creation). This taxonomy is a valuable guide for designing questions that foster deep learning and critical thinking. Using Bloom's revised taxonomy to code learning objectives and test items helps identify and organize cognitive levels. This approach ensures that assessments cover a range of cognitive complexities, from basic recall to higher-order thinking skills.[29]

Research emphasizes the importance of incorporating higher cognitive levels in assessments. For example, a study on quantitative literacy courses showed how different assessment types could evaluate various cognitive levels, ensuring a comprehensive assessment of student proficiency.[30] Additionally, cognitive validation processes, such as those used in performance assessments, are essential for measuring high-level cognitive competencies.[31] As widely recognized, the skill to ask questions is important for teachers. When teachers exclusively pose questions to test students' rote memorization or comprehension, students become adept solely at memorizing, unable to solve problems or interpret information based on their acquired knowledge.

---

[29] T Abate dan E Mishore, "Alignment analysis between teacher-made tests with the learning objectives in a selected school of central regional state of Ethiopia," *Heliyon* 10, no. 11 (2024), https://doi.org/10.1016/j.heliyon.2024.e31869.

[30] Y Bae dkk., "Alignment Between Learning Objectives and Assessments in a Quantitative Literacy Course," *Numeracy* 12, no. 2 (2019), https://doi.org/10.5038/1936-4660.12.2.10.

[31] J P Leighton, "The risk–return trade-off: Performance assessments and cognitive validation of inferences," *British Journal of Educational Psychology* 89, no. 3 (2019): 441–55, https://doi.org/10.1111/bjep.12271.

This study reveals that students cannot devise questions assessing higher cognitive abilities. Evaluation of portfolio tasks indicates that students primarily formulated questions at cognitive levels 3 and 4, focusing on application and analysis. The findings of this study closely align with prior research conducted on prospective biology teachers at Riau University. This earlier research highlighted that these prospective educators exhibited strong proficiency in formulating questions targeting the cognitive domains of C1 (remembering), C2 (understanding), and C3 (applying).[32] The consistency in these results underscores the ability of the prospective teacher to construct assessment items effectively within the foundational levels of Bloom's Taxonomy, reflecting their preparedness to design instructional materials that cater to essential cognitive skills.

Furthermore, the assessment of students' portfolio tasks indicates that nearly all students did not face challenges in developing questions at lower cognitive levels, including abilities such as recalling factual information or concepts, explaining, summarizing, describing, and others. These findings align with other research on prospective elementary school teachers in 2022.[33] According to Bloom's taxonomy, cognitive level 3 is the application, which refers to the ability to use information in new situations and solve problems.[34] Students who can formulate questions at this cognitive level study subjects like religious beliefs and ethics (*Akidah Akhlak*) and Islamic jurisprudence (*Fiqh*). For example, in Aqidah Akhlak, they can formulate problem-solving questions such as: How would you react if you were in that situation?

Meanwhile, cognitive level 4 is breaking down information into components and understanding its structure. The ability to analyze is also often demonstrated by subjects *Akidah Akhlak* and *Fiqh*, such as analyzing someone's behavior, determining the validity of someone's worship if there are specific situations and conditions, and others. However, in other subjects like Islamic cultural history and Qur'an Hadith, there remains a limited presence of questions designed to assess analytical skills.

Several strategies exist for formulating questions that stimulate high-order thinking skills. Firstly, questions should prompt students to apply, analyze, and evaluate

---

[32] Idris, Ferazona, dan Safitri, "Profile of the ability of prospective Biology teachers in making question instruments using Bloom's Taxonomy"; Ricky Lam, *Portfolio Assessment for the Teaching and Learning of Writing*, 2018.

[33] Umar dan Arif Widodo, "Analisis Penggunaan Level Kognitif Dalam Penyusunan Soal Pada Calon Guru Sekolah Dasar," vol. 4, 2022.

[34] B S Bloom dkk., *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain* (New York: David McKay Company, 1956).

information rather than simply recall facts.[35] Secondly, employing real-world problems and case studies encourages students to apply their knowledge in novel situations, fostering more profound understanding and critical thinking.[36] Lastly, open-ended questions with multiple possible answers prompt students to think deeply and articulate their reasoning.[37] These approaches promote robust intellectual engagement and skill development among learners.[38]

Clarity and precision of question

Students' ability to design questions that meet the criteria of clarity and precision reflects their ability to formulate purposeful, unambiguous, and well-structured inquiries that target specific knowledge or skills. Clarity ensures that questions are worded in a straightforward, understandable manner, free of vagueness or ambiguity, enabling respondents to identify the intent without confusion. On the other hand, precision requires that questions are narrowly focused, avoiding overly broad or generic wording and instead focusing on key details, concepts, or applications relevant to the learning objectives. Here are the results of the portfolio assessment:

Table 4. Portfolio Assessment Score

Criteria	Cycle 1		Cycle 2		Cycle 2	
	Score	Category	Score	Category	Score	Category
Clarity and precision of question	55,36	Poor	76,20	Good	82,14	Very Good

Source: Personal documentation, 2024

---

[35] Driana dkk., "Assessing Students' Higher-Order Thinking Skills: Knowledge and Practices of Chemistry Teachers in Vocational Senior Secondary Schools."

[36] I A N T Widhiyani, I N Sukajaya, dan G Suweken, "Pengembangan Soal Higher Order Thinking Skills Untuk Pengkategorian Kemampuan Pemecahan Masalah Geometri Siswa SMP," *Jurnal Pendidikan dan Pembelajaran Matematika Indonesia* 8, no. 2 (2019); Izci dan Caliskan, "Development of prospective teachers' conceptions of assessment and choices of assessment tasks"; Hülya Sönmez, "Prospective Teachers' Activity Designing Skills in Accordance with Cognitive Constructivism Strategies," *Journal of Qualitative Research in Education* 7, no. 4 (2019): 1–24, https://doi.org/10.14689/issn.2148-2624.1.7c.4s.16m; Prima Crismono, "Pengaruh Outdoor Learning Terhadap Kemampuan Berpikir Kritis Matematis Siswa," 2017.

[37] Mochammad Yasir dan Lusiana Dwiyanti, "Analysis of Critical Thinking Levels of Prospective Science Teachers in Ethnoscience Learning Based on Reflective and Impulsive Cognitive Styles," *Jurnal Pendidikan Sains Indonesia* 11, no. 4 (2023): 808–25, https://doi.org/10.24815/jpsi.v11i4.31799; S M Brookhart, *How to Assess Higher-order Thinking Skills in Your Classroom*, Gale virtual reference library (ASCD, 2010).

[38] Ishak Kozikoğlu, "Prospective teachers' cognitive constructs concerning ideal teacher qualifications: A phenomenological analysis based on repertory grid technique," *International Journal of Instruction* 10, no. 3 (2017): 63–78, https://doi.org/10.12973/iji.2017.1035a.

Table 4 shows that in the first cycle, the clarity and precision of the questions were relatively low, with a score of 55.36, which falls into the "Poor" category. This indicates that the questions were likely unclear, vague, or lacked the necessary detail to effectively assess the intended cognitive skills. On the other hand, there was a significant improvement in the clarity and precision of the questions. The score increased to 76.20, moving the category to "Good." This suggests that the questions became clearer, specific, and well-structured, allowing a better assessment of students' understanding and skills.

In the third cycle, the clarity and precision of the questions continued to improve, reaching a score of 82.14, which is categorized as "Very Good." This indicates the questions were clear, precise, and effectively designed to assess the desired cognitive abilities. The consistent improvement across cycles highlights the effectiveness of the interventions or teaching strategies to enhance question formulation.

The clarity and precision of assessment questions significantly impact students' ability to understand and respond effectively. A research article by McTighe and Wiggins discusses the importance of clear and precise language in assessment items. They argue that ambiguous wording can lead to misinterpretation and frustration among students, ultimately affecting their performance. The authors recommend using straightforward language and providing clear instructions to enhance student understanding.[39] In a study, Sadler explores the role of clarity in assessment criteria. The findings suggest that students are better equipped to meet expectations when articulating assessment criteria. The study emphasizes the need for educators to provide explicit guidelines and examples to clarify assessment requirements.[40]

The clarity and precision of assessment questions are critical for valid and reliable measurement. The AER includes criteria for evaluating the clarity of assessment instructions and results, ensuring they are interpretable and unambiguous. The AER is designed to guide the development and evaluation of assessments, focusing on four key features: alignment with learning goals, the formative or summative nature of the appraisal, cognitive complexity, and clarity of the evaluation.[41] Another study discussed the importance of clear and meaningful learning objectives, which directly impact the

---

[39] Jay B Y Mctighe dan Grant Wiggins, "Understanding By Design ® Framework Introduction: What Is Ubd^TM Framework?," 2016.

[40] D R Sadler, "Formative Assessment and the Design of Instructional Systems," *Instructional Sciences* 45, no. 1 (2017): 1–24.

[41] Tractenberg, "The assessment evaluation rubric: Promoting learning and learner-centered teaching through assessment in face-to-face or distanced higher education."

formulation of precise assessment questions.[42] Various procedures for evaluating survey questions consider aspects such as complexity, precision, and potential biases. These procedures help refine questions to ensure they accurately measure the intended concepts.[43]

In this study, the prevalent errors identified in the questions include incorrect grammar, such as incomplete sentence structures, placing question words at the end of sentences, and mixing commands with questions in a single sentence; Usage of non-standard or inconsistent terminology and vocabulary; Improper use of negative constructions, illustrated by phrases like: "The following is not included in the pillars of prayer, except…" Several strategies for formulating clear and precise questions measure what is intended to be assessed, as follows: Using simple, direct language ensures that questions are easily understood.[44] Questions should be specific enough to guide students on what is expected without providing too much information.[45] Consistent formatting and terminology throughout an assessment can help maintain clarity and precision.[46]

Portfolio-based learning tasks encourage students to independently strive to understand and practice the skills they are learning. According to this research, several crucial factors that enhance the effectiveness of portfolio tasks are as follows: Giving an example for each assignment or thematic task, the teacher provides an example for every assigned task, requesting students to complete 50% of their tasks during ongoing learning. This ensures that students understand what is required to fulfill the assigned task. After each student presents their work, the teacher provides feedback on whether it meets the task requirements. During each cycle, the researcher allocates 40-50 minutes for students to consult on the tasks they are working on, offering feedback on each student's task. Reviewing and providing feedback on students' tasks is consistently conducted at the end of each cycle. This practice is crucial for students to recognize

---

[42] A A Sewagegn, "Learning objective and assessment linkage: Its contribution to meaningful student learning," *Universal Journal of Educational Research* 8, no. 11 (2020): 5044–52, https://doi.org/10.13189/ujer.2020.081104.

[43] A J Mahamuni dan S S Tonpe, "Enhancing Educational Assessment with Artificial Intelligence: Challenges and Opportunities," dalam *2024 International Conference on Knowledge Engineering and Communication Systems, ICKECS 2024* (Institute of Electrical and Electronics Engineers Inc., 2024), https://doi.org/10.1109/ICKECS61492.2024.10616620.

[44] Sönmez, "Prospective Teachers' Activity Designing Skills in Accordance with Cognitive Constructivism Strategies"; Susanne Narciss dkk., "Promoting Prospective Teacher Competencies for Designing, Implementing, Evaluating, and Adapting Interactive Formative Feedback Strategies," *Psychology Learning and Teaching* 20, no. 2 (2020): 261–78, https://doi.org/10.1177/1475725720971887.

[45] W J Popham, *Classroom assessment: What teachers need to know* (Pearson, 2010).

[46] M D Miller, R L Linn, dan N E Gronlund, *Measurement and Assessment in Teaching* (Pearson, 2009).

errors in their completed tasks and improve their performance on subsequent assignments.

Conclusion

The findings of this study confirm the pedagogical value of portfolio assignments in developing students' test design skills, particularly their ability to formulate questions that align with learning objectives, measure higher-order cognitive skills, and prioritize clarity and precision. By involving students in iterative assessment, evaluation, and refinement processes, portfolios foster a deeper understanding of constructive alignment - ensuring that assessment is a deliberate extension of curriculum goals rather than a mere administrative exercise. This approach emphasizes the symbiotic relationship between assessment design and pedagogical intent. It positions students as creators and critical evaluators of tools that produce meaningful learning outcomes.

These findings have important implications for educational practice. Institutions training future educators should integrate portfolio-based methods into their curricula to bridge the gap between theoretical knowledge and practical assessment skills. Such training can empower educators to design assessments beyond memorization to emphasize analytical thinking and problem-solving. In addition, professional development programs for in-service teachers should emphasize reflective practices. This will encourage educators to critically evaluate their assessment philosophies and adopt frameworks that balance cognitive rigor and equity. The study also highlights the need for interdisciplinary collaboration between curriculum designers, assessment specialists, and policymakers to standardize adaptable, HOTS-oriented assessment frameworks across disciplines.

Future studies should explore the longitudinal impact of portfolio-based training on educators' assessment practices and student learning outcomes to build on these insights. Comparative research could investigate the efficacy of portfolios versus traditional assessment-design training models, particularly in diverse cultural or institutional contexts. Further examination is needed into how technology-enhanced portfolios—such as digital platforms with AI-driven feedback—might streamline iterative refinement processes or address scalability challenges. Additionally, research should probe the role of socio-cultural factors in shaping educators' adoption of portfolio methodologies, including barriers related to resource availability, institutional

support, or entrenched pedagogical traditions. Finally, interdisciplinary studies could explore the applicability of portfolio frameworks in non-educational sectors, such as corporate training or public policy, where assessment literacy is increasingly critical.

Acknowledgment

References

Abate, T, dan E Mishore. "Alignment analysis between teacher-made tests with the learning objectives in a selected school of central regional state of Ethiopia." *Heliyon* 10, no. 11 (2024). https://doi.org/10.1016/j.heliyon.2024.e31869.

Adam Smith, R, dan L Robinson. "The impact of clear scoring criteria on student performance: A systematic review. Journal of Educational Assessment." *Journal of Educational Assessment* 12, no. 3 (2024): 45–67.

Ayu, Betzy, Omega Rampean, Eli Rohaeti, dan Widya Prasasti Utami. "Teacher Difficulties for Develop Higher Order Thinking Skills Assessment Instrument on Reaction Rate." *Jurnal Pendidikan Kimia Indonesia* 6 (2022): 11–19. https://doi.org/10.23887/jpk.v6i1.

Bae, Y, S L Tunstall, K S Knowles, dan R L Matz. "Alignment Between Learning Objectives and Assessments in a Quantitative Literacy Course." *Numeracy* 12, no. 2 (2019). https://doi.org/10.5038/1936-4660.12.2.10.

Bloom, B S, M D Engelhart, E J Furst, W H Hill, dan D R Krathwohl. *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York: David McKay Company, 1956.

Brookhart, S M. *How to Assess Higher-order Thinking Skills in Your Classroom*. Gale virtual reference library. ASCD, 2010.

Cizek, G J, A E Kosh, dan E K Toutkoushian. "Gathering and Evaluating Validity Evidence: The Generalized Assessment Alignment Tool." *Journal of Educational Measurement* 55, no. 4 (2018): 477–512. https://doi.org/10.1111/jedm.12189.

Crismono, Prima. "Pengaruh Outdoor Learning Terhadap Kemampuan Berpikir Kritis Matematis Siswa," 2017.

———. "The Effect of Outdoor Learning on Students' Attitudes in Mathematics Learning." *Jurnal Axioma: Jurnal Matematika dan Pembelajaran* 8, no. 2 (2023).

Crismono, Prima Cristi. *Metode Penelitian Pendidikan : Pendekatan Kuantitatif, Kualitatif, Ptk, dan Penelitian Pengembangan*. Disunting oleh Fitriyatul Hanifiyah dan Iqbal Erdiansyah. Bondowoso: KHD Production, 2024.

———. "Pengaruh Outdoor Mathematich Terhadap Kemampuan Berpikir Tingkat Tinggi (HOTS)." *FAJAR Jurnal Pendidikan Islam* 3, no. 2 (2023): 160–67.

Crismono, Prima Cristi, Christine Wulandari Suryaningrum, dan Tri Endang Jatmikowati. "Pengaruh model pembelajaran connecting, organizing, reflecting, extending terhadap hasil belajar siswa pada pembelajaran matematika." *SIGMA* 9, no. 2 (2024): 135–40.

Daher, Wajeeh, Nimer Baya'a, dan Otman Jaber. "Understanding Prospective Teachers' Task Design Considerations through the Lens of the Theory of Didactical Situations." *Mathematics* 10, no. 3 (2022). https://doi.org/10.3390/math10030417.

Driana, Elin, Ary Susilowati, Ernawati Ernawati, dan Abd. Rahman A Ghani. "Assessing Students' Higher-Order Thinking Skills: Knowledge and Practices of Chemistry Teachers in Vocational Senior Secondary Schools." *Jurnal Pendidikan Teknologi dan Kejuruan* 27, no. 1 (2021): 37–47. https://doi.org/10.21831/jptk.v27i1.32882.

Hidayati, Nurkhairo, Tengku Idris, Peny Husna Handayani, dan Fitri Arsih. "Portfolio assessment with dimension of learning: an approach on the mastery of concept." *JURNAL BIOEDUKATIKA* 9, no. 1 (2021): 17. https://doi.org/10.26555/bioedukatika.v9i1.19384.

Idris, Tengku, Sepita Ferazona, dan Herlina Safitri. "Profile of the ability of prospective Biology teachers in making question instruments using Bloom's Taxonomy." *REID (Research and Evaluation in Education)* 7, no. 2 (2021): 177–85. https://doi.org/10.21831/reid.v7i2.44903.

Izci, Kemal, dan Gurbuz Caliskan. "Development of prospective teachers' conceptions of assessment and choices of assessment tasks." *International Journal of Research in Education and Science* 3, no. 2 (2017): 464–74. https://doi.org/10.21890/ijres.327906.

Johnson, M, dan T Lee. "The role of clear learning outcomes in student engagement and performance." *Educational Research Review* 16, no. 2 (2021): 78–92.

Kozikoğlu, Ishak. "Prospective teachers' cognitive constructs concerning ideal teacher qualifications: A phenomenological analysis based on repertory grid technique." *International Journal of Instruction* 10, no. 3 (2017): 63–78. https://doi.org/10.12973/iji.2017.1035a.

Lam, Ricky. "E-Portfolios for self-regulated and co-regulated learning: A review." *Frontiers in Psychology*. Frontiers Media S.A., 2022. https://doi.org/10.3389/fpsyg.2022.1079385.

———. *Portfolio Assessment for the Teaching and Learning of Writing*, 2018.

Lee Kim, J., S, dan H Park. "Higher-order thinking in assessments: A comparative study." *International Journal of Educational Research* 112 (2023): 101–15.

Leighton, J P. "The risk–return trade-off: Performance assessments and cognitive validation of inferences." *British Journal of Educational Psychology* 89, no. 3 (2019): 441–55. https://doi.org/10.1111/bjep.12271.

Linanti, Tiara Annisa, Saiful Ridlo, dan Siti Harnina Bintari. "The Implementation of Portfolio Assessment to Increase Critical Thinking Ability for High School

Students on Human Coordination System Material." *Journal of Innovative Science Education* 10, no. 2 (2021): 130–36.

Mahamuni, A J, dan S S Tonpe. "Enhancing Educational Assessment with Artificial Intelligence: Challenges and Opportunities." Dalam *2024 International Conference on Knowledge Engineering and Communication Systems, ICKECS 2024*. Institute of Electrical and Electronics Engineers Inc., 2024. https://doi.org/10.1109/ICKECS61492.2024.10616620.

Martinez Lopez, R., A, dan M Garcia. "Validity in educational assessments: Content and construct considerations." *Educational Measurement: Issues and Practice* 42, no. 1 (2023): 15–29.

Mctighe, Jay B Y, dan Grant Wiggins. "Understanding By Design ® Framework Introduction: What Is Ubd^{TM} Framework?," 2016.

Miller, M D, R L Linn, dan N E Gronlund. *Measurement and Assessment in Teaching*. Pearson, 2009.

Murphy, Joseph, Ellen Goldring, Xiu Cravens, dan Stephen Nelson Elliott. "The Vanderbilt Assessment of Leadership in Education: Measuring Learning-Centered Leadership," 2014.

Narciss, Susanne, Elsa Hammer, Gregor Damnik, Kerstin Kisielski, dan Hermann Körndle. "Promoting Prospective Teacher Competencies for Designing, Implementing, Evaluating, and Adapting Interactive Formative Feedback Strategies." *Psychology Learning and Teaching* 20, no. 2 (2020): 261–78. https://doi.org/10.1177/1475725720971887.

Nguyen, T, dan L Tran. "Integrating Bloom's Taxonomy in assessment design: Fostering higher-order thinking." *Journal of Curriculum Studies* 57, no. 2 (2025): 134–50.

Patel, R, dan S Kumar. "The importance of test blueprints in educational assessment." *Assessment & Evaluation in Higher Education* 49, no. 4 (2023): 567–80.

Popham, W J. *Classroom assessment: What teachers need to know*. Pearson, 2010.

Prihandoko, Yogi, dan Elementary School Teacher Training. "International Journal of Curriculum Development, Teaching and Learning Innovation The effectiveness of portfolio assessment based problem based learning on mathematical critical thinking skills in elementary schools," 2023.

Rahmawati, T, T Mulyaningsih, N Nahadi, H Suhanda, W K Lee, H A Aziz, dan S Anwar. "Electronic Portfolio Assessment Instruments In Improving Students' Creative Thinking Skills." *Jurnal Pendidikan IPA Indonesia* 12, no. 4 (2023): 598–610. https://doi.org/10.15294/jpii.v12i4.45639.

Rizal, M Khairu, Dwi Yulianti, Handoko, dan Udang Rosidin. "Portfolio Assessment to Measuring Capabilities Communications and Collaborations of Grade V Students in Elementary School." *Journal of Education Action Research* 7, no. 4 (2023): 559–68. https://doi.org/10.23887/jear.v7i4.70112.

Robinson, L, dan J Smith. "The effects of clear instructions on student assessment performance." *Journal of Educational Psychology* 114, no. 3 (2022): 345–60.

Sadler, D R. "Formative Assessment and the Design of Instructional Systems." *Instructional Sciences* 45, no. 1 (2017): 1–24.

Sewagegn, A A. "Learning objective and assessment linkage: Its contribution to meaningful student learning." *Universal Journal of Educational Research* 8, no. 11 (2020): 5044–52. https://doi.org/10.13189/ujer.2020.081104.

Sönmez, Hülya. "Prospective Teachers' Activity Designing Skills in Accordance with Cognitive Constructivism Strategies." *Journal of Qualitative Research in Education* 7, no. 4 (2019): 1–24. https://doi.org/10.14689/issn.2148-2624.1.7c.4s.16m.

Thompson, G, dan M Garcia. "The impact of assessment alignment on student achievement: A meta-analysis." *Educational Psychology Review* 35, no. 2 (2023): 215–30.

Tractenberg, R E. "The assessment evaluation rubric: Promoting learning and learner-centered teaching through assessment in face-to-face or distanced higher education." *Education Sciences* 11, no. 8 (2021). https://doi.org/10.3390/educsci11080441.

Umar, dan Arif Widodo. "Analisis Penggunaan Level Kognitif Dalam Penyusunan Soal Pada Calon Guru Sekolah Dasar." Vol. 4, 2022.

White, Parker C. "Student portfolios: An alternative way of encouraging and evaluating student learning." *New Directions for Teaching and Learning* 2004, no. 100 (2004): 37–42. https://doi.org/10.1002/tl.169.

Widhiyani, I A N T, I N Sukajaya, dan G Suweken. "Pengembangan Soal Higher Order Thinking Skills Untuk Pengkategorian Kemampuan Pemecahan Masalah Geometri Siswa SMP." *Jurnal Pendidikan dan Pembelajaran Matematika Indonesia* 8, no. 2 (2019).

Wulandari, Putri Maulidah, Prima Cristi Crismono, dan Muhammad Ilyas. "Pengaruh Aplikasi Media Papan Pintar Terhadap Kemampuan Berpikir Kreatif Siswa Pembelajaran Tematik." *Jurnal Pendidikan Matematika dan Sains* 12, no. 2 (2024). https://doi.org/10.21831/jpms.v12i2.72230.

Yanuardianto, Elga, Djoko Rohadi Wibowo, dan Prima Cristi Crismono. "Innovation in Islamic Religious Education in The Digital Era for Students' Religious Tolerance Character Building in Elementary School: An Implementation of Beyond the Wall Learning Model." *JIP (Jurnal Ilmiah PGMI)* 10, no. 2 (2024): 49–61.

Yasir, Mochammad, dan Lusiana Dwiyanti. "Analysis of Critical Thinking Levels of Prospective Science Teachers in Ethnoscience Learning Based on Reflective and Impulsive Cognitive Styles." *Jurnal Pendidikan Sains Indonesia* 11, no. 4 (2023): 808–25. https://doi.org/10.24815/jpsi.v11i4.31799.

Zhang, Chen X., dan H Liu. "Aligning assessment methods with learning objectives: A systematic approach." *Assessment in Education: Principles, Policy & Practice* 29, no. 1 (2022): 23–40.