TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC SƯ PHẠM TP HỒ CHÍ MINH
Tập 22, Số 9 (2025): 1619-1630
Website: https://journal.hcmue.edu.vn
ISSN: 2734-9918

HO CHI MINH CITY UNIVERSITY OF EDUCATION
JOURNAL OF SCIENCE
Vol. 22, No. 9 (2025): 1619-1630
https://doi.org/10.54607/hcmue.js.22.9.4706(2025)

Research Article

DEVELOPING A MULTIPLE-CHOICE QUESTION BANK TO ASSESS LEARNING OUTCOMES FOR THE COURSE 'FOUNDATIONS OF MATHEMATICS IN ELEMENTARY EDUCATION 1' AT HO CHI MINH CITY UNIVERSITY OF EDUCATION

Nguyen Viet Khoa**, Phan Le

Ho Chi Minh City University of Education, Vietnam

**Corresponding author: Nguyen Viet Khoa – Email: khoanvi@hcmue.edu.vn*

Received: January 21, 2025; Revised: March 14, 2025; Accepted: April 18, 2025

ABSTRACT

Assessing student learning outcomes is a crucial component of university education, playing a vital role in ensuring and enhancing educational quality. This research focuses on developing a multiple-choice question (MCQ) bank for the course 'Foundations of Mathematics in Elementary Education 1' to support student assessment. The study employs a mixed-methods approach, combining qualitative and quantitative methods, and draws upon theoretical foundations of test construction and MCQ design. The question bank development follows a rigorous, five-step scientific process. A total of 300 multiple-choice questions were initially developed, of which 30 were piloted on a sample of 140 students. The pilot results confirmed the feasibility and initial effectiveness of the question bank, while also highlighting areas for improvement. Specifically, 13.3% of the items were classified as difficult, 23.3% as moderate, and 63.3% as easy, based on item difficulty indices. The overall test difficulty was 0.7113—slightly above the optimal range (0.555-0.695)—indicating the test was relatively easy for the target group. These results suggest a need to adjust the distribution of item difficulty, particularly by increasing the proportion of moderate and difficult questions, improving content clarity, and broadening the sample in future pilots. Such adjustments aim to enhance the test's discriminatory capacity, reliability, and alignment with course learning outcomes.

Keywords: learning outcomes assessment; mathematics education; multiple-choice question bank; objective testing

1. Introduction

Assessing and evaluating learning outcomes are fundamental aspects of university education, ensuring and enhancing instructional quality (Dang et al., 2024). In Vietnam's educational reform, assessment not only measures student knowledge but also informs

Cite this article as: Nguyen, V. K., & Phan, L. (2025). Developing a multiple-choice question bank to assess learning outcomes for the course 'foundations of Mathematics in elementary education 1' at Ho Chi Minh City University of Education. *Ho Chi Minh City University of Education Journal of Science, 22*(9), 1619-1630. https://doi.org/10.54607/hcmue.js.22.9.4706(2025)

teaching improvements and supports student self-assessment. The Central Executive Committee's Resolution (2013) emphasizes shifting education from knowledge acquisition to developing student competencies and qualities. Accurate assessment aids instructors in refining teaching methods, enhances student learning, and optimizes educational management (Ho et al., 2013).

Globally, multiple-choice questions (MCQs) are widely used due to their objectivity, reliability, and efficiency in assessing various competencies within a limited timeframe (Faniyi, 2023; Jia & Zhu, 2020; Liu et al., 2023). While some critique MCQs for emphasizing memorization over deeper understanding, they remain essential in standardized testing, offering broad coverage and reducing grading bias (Bailey et al., 2012; DiBattista & Kurzawa, 2011). Despite requiring careful preparation, MCQs provide accurate competency assessment (Tran & Ngo, 2024; Vo et al., 2013).

Since transitioning to a credit-based system in 2010, Ho Chi Minh City University of Education (HCMUE) has encouraged objective testing to enhance educational quality (Ho et al., 2013). To improve assessment in Foundations of Mathematics in Elementary Education 1, developing and piloting an MCQ bank is essential. This study addresses two key questions: (1) What steps are involved in developing an MCQ bank for this course? and (2) What are the results of the pilot test?

2. Research Content

2.1. Relevant Concepts

Objective testing: Objective testing is a systematic process or tool used to measure an individual's abilities for specific purposes (Hoang, 2008). It is also defined as a structured procedure that enables individuals to determine their level of achievement in a specific area (Tran, 2007). Pre-set answer options for each question characterize objective tests. This format provides some or all of the necessary information and requires students to select an answer or fill in a few words, also known as closed-ended questions. Objective tests are considered objective because the grader does not influence the scoring system. Objective tests must be designed to ensure accuracy so that each question has only one correct or "best" answer, typically requiring a simple indication (Dang, 2003).

Course Learning Outcomes: These are the expected achievements in terms of qualities and competencies that students should attain after completing a course, including minimum requirements for knowledge, skills, self-directed learning, and responsibility. These outcomes must be clearly and comprehensively defined, measurable, and assessable across different cognitive levels and serve as a basis for designing, implementing, and improving content and teaching methods and evaluating learning. The Ministry of Education and Training (MOET, 2021a, 2021b) emphasizes that programs should be coherently organized through interrelated course components to guarantee the attainment of program learning outcomes.

Characteristics of the 'Foundations of Mathematics in Elementary Education 1'

This is a compulsory foundational course divided into five chapters. Through this course, students understand set theory concepts, the rules for constructing basic number sets (natural numbers, integers, rational numbers, and real numbers), and operations on these sets. Students can also explain the theory of divisibility and remainders on integers and apply the theory of congruences to solve related real-world problems. They can also define and apply knowledge of prime numbers, divisors, and multiples in practical situations. Based on this foundation, students can select appropriate basic algorithms and mathematical reasoning when solving practical problems, practice thinking and mathematical reasoning, cultivate critical thinking, and contribute to problem-solving and creative abilities (HCMUE, 2024).

Assessment Methods for 'Foundations of Mathematics in Elementary Education 1'

Assessment methods for this course must be based on course learning outcomes, clarifying the level of student achievement according to cognitive levels defined in the course outcomes. Assessment should be based on both formative and summative evaluations, providing a basis for timely adjustments to teaching and learning activities, motivating and supporting student progress, improving the curriculum, and implementing the training program.

2.2. Research Methods

This research reviews relevant literature on objective assessment for learning outcomes. Additionally, the study also used a multiple-choice test for 'Foundations of Mathematics in Elementary Education 1.'

2.3. Process of Developing the Multiple-Choice Question Bank for 'Foundations of Mathematics in Elementary Education 1'

2.3.1. Principles of Question Bank Development

Based on the above analysis, we identify the basis for developing questions to assess learning outcomes as follows: design program learning outcomes; content and requirements of each teaching topic; meaningful situations and real-world problems related to the content of the teaching topics; test structure (HCMUE, 2024; MOET, 2021). The development of MCQs for assessing student learning outcomes follows these principles: 1) Ensuring alignment with objectives; 2) Ensuring scientific rigor; 3) Ensuring reliability and validity; 4) Ensuring appropriateness; and 5) Ensuring feasibility (Dang, 2003; Dang et al., 2024; Ho et al., 2013; Vo et al., 2013; Tran, 2007).

2.3.2. Steps in Developing the Question Bank

A scientific process is essential for creating a standardized MCQ bank, ensuring validity and reliability in assessment. The development process follows a structured sequence (Dang et al., 2024; Ho et al., 2013; Vo et al., 2013), where each step builds upon the previous one to achieve a well-constructed test. Step 1 (Determining objectives and content) establishes the foundation by defining what the test should measure, guiding Step 2 (Designing the Test Blueprint) to organize question distribution effectively. The number and levels of questions in this step are not arbitrary; rather, they are based on the structure of the course, which includes five chapters of relatively equal weight, and are aligned with four cognitive levels of Bloom's taxonomy. This proportional distribution ensures comprehensive coverage of both the course content and learning outcomes, while maintaining a balance in assessing lower- and higher-order thinking skills (HCMUE, 2024; MOET, 2021a). Step 3 (Drafting, expert consultation, and revision) then enhances item quality and alignment. Step 4 (Selecting a student sample for piloting) validates the test through real student responses, allowing Step 5 (Scoring and item analysis using IATA software) to refine and finalize the test based on statistical analysis.

Step 1. Determine the objectives of the test and the main content of each section in the course

The main content of each section is shown in Table 1.

Table 1. *Objectives and main content of 'Basic Mathematics in Primary School 1'*

Chapters	Learning Objectives (Knowledge-based)	Main Content Structure
Chapter 1: Sets	- Present the results of operations on sets - Explain Venn diagrams and determine the cardinality of finite sets - Apply set theory to solve specific problems	- Operations on sets - Cardinality of a finite set - Some problems in set theory
Chapter 2: Binary Relations	- Present the concepts of equivalence relations and order relations - Explain the properties of equivalence relations and the partition of equivalence classes - Apply order relation theory to find the smallest and largest elements	- Binary relations and their properties - Equivalence relations and the partition of equivalence classes - Order relations and special elements in a set
Chapter 3: Natural Numbers	- Present various numeral systems - Describe models of natural numbers - Apply mental calculation methods to solve problems	- Formation of the set of natural numbers - Models of natural numbers - Mental calculation methods in the set of natural numbers
Chapter 4: Integers and Rational Numbers	- Present models of integers - Explain the representation of fractions - Apply theory to solve problems involving ratios and percentages	- Integers and models of integers - Fractions and their representation - Decimal numbers, ratios, and percentages
Chapter 5: Number Theory	- Present the concept of the Euclidean algorithm - Explain the properties of congruence theory - Solve systems of first-degree congruence equations with one variable - Apply Euler's theorem to determine remainders in division problems	- Euclidean algorithm, greatest common divisor, and least common multiple - Congruences, properties of congruences, Euler's theorem, and applications - Equations and systems of first-degree congruence equations with one variable

Step 2. Designing the Test Blueprint

For the *Foundations of Mathematics in Elementary Education 1* course, the test blueprint is designed using a two-dimensional matrix to ensure a structured and comprehensive assessment. One dimension represents the content knowledge of the five course chapters, while the other aligns with the four cognitive levels of Bloom's taxonomy: remembering, understanding, applying, and analyzing. The number of questions at each cognitive level is clearly defined, with a total of 300 MCQs, evenly distributed across the five chapters (60 questions per chapter). The blueprint is finalized before drafting the MCQ items to maintain consistency and alignment with assessment objectives (Table 2).

Table 2. *Blueprint for the 'Foundations of Mathematics in Elementary Education 1' Test*

Chapter	Remembering	Understanding	Applying	Analyzing	Total
Chapter 1: Sets	14	22	18	6	60
Chapter 2: Binary Relations	13	21	19	7	60
Chapter 3: Natural Numbers	15	20	16	9	60
Chapter 4: Integers and Rational Numbers	15	22	18	5	60
Chapter 5: Number Theory	12	20	20	8	60
Total	69	105	91	35	300

Step 3. Drafting, Consulting Experts, and Preliminary Revision of Questions based on the Two-Dimensional Blueprint for each Content Area

Based on the course objectives, the content of each section, and the two-dimensional blueprint for both the question bank and each test, the research team assigned members to write and select questions appropriate for each section's knowledge scope and testing range. The team used single-choice questions for the test. During the development process, members adhered to the two-dimensional blueprint and followed basic principles, including ensuring scientific rigor in the drafting process, aligning with student understanding, and balancing the accuracy and feasibility of each answer option. After completing the question bank, the team conducted a preliminary analysis and sought expert opinions to promptly identify and address any errors. Issues such as question phrasing, student test-taking time, and distractor effectiveness were identified. Based on the feedback, the team members reviewed, revised, and finalized the question bank before piloting.

Step 4. Selecting a Student Sample for Piloting the Objective Test Questions

The objective of the first trial was to grade, analyze the test, and evaluate the multiple-choice questions based on essential statistical indicators such as mean, standard deviation, difficulty level, and reliability of the test and individual questions. Based on the discrimination index of each question, the answers and distractors were analyzed to decide which questions with a satisfactory discrimination index or higher should be included in the test or question bank. Questions with low discrimination indices were revised and further tested, and analyzed. Due to time constraints, the study focused on testing a 30-question multiple-choice exam with 140 second-year students (17 male, 123 female) from the Primary Education Departments at Ho Chi Minh City University of Education. These students completed the course 'Foundations of Mathematics in Elementary Education 1' in the first semester of the 2024–2025 academic year. The outline of the multiple-choice exam (first trial) is presented in Table 3 as follows:

Table 3. *Outline of the Multiple-Choice Exam for 'Foundations of Mathematics in Elementary Education 1'*

Chapters	Remembering	Understanding	Applying	Analyzing	Total
Chapter 1: Sets	1	2	2	1	6
Chapter 2: Binary Relations	2	2	1	1	6
Chapter 3: Natural Numbers	1	2	1	1	5
Chapter 4: Integers and Rational Numbers	1	2	2	1	6
Chapter 5: Number Theory	2	2	2	1	7
Total	**7**	**10**	**8**	**5**	**30**

Step 5. Scoring, Item Analysis, and Test Analysis using IATA Software

2.3.3. Evaluating the Difficulty and Reliability of the Objective Test Questions

Item Difficulty

Item difficulty refers to how easy or difficult a question is on a test. It is calculated as the proportion of students who answered the question correctly. The formula for calculating item difficulty (P) is:

$$P_{item\ i} = \frac{Number\ of\ correct\ answers\ by\ students}{Total\ answers\ by\ students}$$ (Duong, 1995, p. 64; Lam, 2008, p. 59).

The item difficulty value ranges from 0.0 to 1.0 (corresponding to 0% to 100%). Item difficulty has an inverse relationship with the P-value; as P approaches 0, the question becomes more difficult, and as P approaches 1, the question becomes easier.

Optimal Item Difficulty

Items with optimal difficulty (neither too easy nor too difficult) contribute to the effectiveness of the test. The optimal difficulty level varies depending on the question type. A difficulty of 50% is not considered optimal for True/False questions, as there's a 50% chance of guessing correctly (Duong, 1995). Therefore, when determining optimal difficulty, the guessing factor should be considered to ensure that the difficulty accurately reflects students' actual abilities. For four-option multiple-choice questions, the probability of guessing correctly is 25%. Thus, the optimal difficulty for four-option multiple-choice questions can be calculated as:

$$P_{optimal} = \frac{100\% + Risk\ probability}{2} = \frac{100\% + 25\%}{2} = 62.5\%$$

In this study, all 30 questions are four-option multiple-choice questions. The difficulty of each question is determined by comparing it to the optimal difficulty level of 62.5%. Questions with difficulty exceeding 62.5% are generally considered easy. Conversely, questions with difficulty below 62.5% tend to be difficult. Based on this principle, the appropriate difficulty range for the multiple-choice questions is determined by applying a margin of ±7% around the optimal difficulty. In this case, the acceptable range is from 55.5% to 69.5% (equivalent to 0.555 to 0.695). The results show that there are 4 difficult questions, 7 moderately difficult questions, and 19 easy questions.

Table 4. *Item Difficulty Analysis*

Item	Correct Answer	P-Score	Classification	Item	Correct Answer	P-Score	Classification
C1	107	76.4	Easy	C16	111	79.3	Easy
C2	99	70.7	Easy	C17	83	59.3	Moderate
C3	72	51.4	Difficult	C18	112	80	Easy
C4	68	48.6	Difficult	C19	111	79.3	Easy
C5	102	72.9	Easy	C20	119	85	Easy
C6	65	46.4	Difficult	C21	111	79.3	Easy
C7	87	62.1	Moderate	C22	109	77.9	Easy
C8	102	72.9	Easy	C23	94	67.1	Moderate
C9	88	62.9	Moderate	C24	92	65.7	Moderate
C10	122	87.1	Easy	C25	107	76.4	Easy
C11	114	81.4	Easy	C26	93	66.4	Moderate
C12	111	79.3	Easy	C27	105	75	Easy
C13	99	70.7	Easy	C28	76	54.3	Difficult
C14	117	83.6	Easy	C29	108	77.1	Easy
C15	111	79.3	Easy	C30	93	66.4	Moderate

The analysis of item difficulty in Table 4 reveals an imbalance in question distribution, with a predominance of easy questions and very few moderate or difficult ones. This imbalance reduces the test’s effectiveness in differentiating student abilities, as an excessive number of easy questions may fail to assess critical thinking skills. To improve assessment quality, the proportion of questions should be adjusted to achieve a better balance across difficulty levels. Difficult questions should be reviewed for clarity and presentation to ensure they align with student abilities, while easy questions should incorporate elements that encourage higher-order thinking without compromising accessibility. Additionally, diversifying content coverage is essential to prevent overemphasis on a single topic and ensure a comprehensive evaluation of student knowledge. In line with these improvements, Iñarrairaegui et al. (2022) recommend increasing the use of case-based questions, adjusting question difficulty levels, and eliminating non-functioning distractors to enhance the discriminative power of multiple-choice assessments. Pilot testing on a smaller group before full implementation will help assess question appropriateness and difficulty. Finally, gathering feedback from students and instructors can further refine the test, ensuring it remains a reliable and comprehensive assessment tool.

Test Difficulty

Test difficulty indicates how easy or difficult the entire test is. It is calculated based on the average score of all test-takers compared to the maximum possible score. The formula for test difficulty is:

$$P_{test} = \frac{Average\ score\ of\ students\ (M)}{Max\ score\ of\ the\ test\ (K)} x\ 100\% = \frac{21.34}{30} \approx 71.13\%$$

Optimal Test Difficulty

Optimal test difficulty is achieved when the questions effectively differentiate between high and low-performing students while ensuring the test is neither too easy nor too difficult.

The optimal difficulty for the test is determined when the questions effectively discriminate between high- and low-performing students while ensuring the test is neither too easy nor too difficult. The optimal difficulty for the test is calculated as follows:

$$P_{test\ optimal} = \frac{M_{LT}}{K}$$

Where:

- M_{LT}: Theoretical average score
- T: Guessing score ($T = \frac{K}{number\ of\ answer\ choices} = \frac{30}{4} = 7.5$)
- K: Total number of questions on the test (N = 30)

In this case, $M_{LT} = \frac{K+T}{2} = \frac{30+7.5}{2} = 18.75$. Therefore, the optimal test difficulty is:

$$P_{test\ optimal} = \frac{M_{LT}}{K}\ x\ 100\% = \frac{18.75}{30}\ x\ 100\% = 62.5\%$$

A test is considered appropriate for students when its difficulty falls within a range defined by the optimal difficulty ± 0.07. In this case, the range is 62.5% or 0.625 ± 0.07. If the test difficulty is below the optimal difficulty, the test is considered easy. Conversely, if the test difficulty exceeds the optimal difficulty, the test is considered difficult for students. The results of this study indicate that the actual test difficulty ($P_{test} = 71.13\%$) falls outside the upper limit of the optimal difficulty range (55.5% to 69.5%). Therefore, the test is considered slightly easier for the students.

Table 5. *Results of the Pilot Test Analysis (First Trial)*

Number of Tests	Average Score	Theoretical Average	Standard Deviation	Test Difficulty	Optimal Difficulty	Reliability (KR-21)
140	21.34	18.75	3.26	0.7113	0.625	0.463

The pilot test has an average score (21.34) higher than the theoretical average (18.75), suggesting the test might be slightly easier for the participants. This aligns with the test difficulty (0.7113) and optimal difficulty (0.625) values, indicating that most questions are easy or moderate. The standard deviation (3.26) is relatively low, indicating that the scores are not widely dispersed, meaning there is not a significant difference between high and low scores. This also suggests that the test does not effectively discriminate among students with varying abilities. The average difficulty of the test falls within the easy range (0.7113), as most students answered correctly, but it may lack challenging questions, limiting its ability to differentiate students. The proportion of "moderate" difficulty questions (0.625) is reasonable but could be increased to improve discrimination. The reliability coefficient (KR-21 = 0.463) is lower than the acceptable standard of 0.7, indicating the test has low reliability, potentially due to inconsistencies in question content or inadequate coverage of the assessed skills.

To improve the test, the difficulty level needs to be balanced by increasing the proportion of moderate and difficult questions to reflect students' abilities at different levels accurately. The content should be revised and supplemented with application-oriented questions and those requiring higher-order thinking skills, while also strengthening the interrelation between questions to ensure the test aligns with one or more specific learning objectives. Reviewing the test structure, removing inappropriate questions, and piloting the revised version on different sample groups are also necessary to improve reliability and assess test stability. Overall, the current test is suitable for a large portion of the student population, but the lack of discrimination and low reliability need to be addressed to ensure the test meets the standards of objectivity and effectiveness.

Test Reliability (Kuder-Richardson Formula 21 - KR-21)

The Kuder-Richardson Formula 21 (KR-21) is used to measure the reliability of the test (Duong, 2005, p. 150; Lam, 2008, pp. 69-71). This is a quick and simple way to assess the internal consistency of the test. The KR-21 formula is:

$$KR_{21} = \frac{K}{K-1}\left(1-\frac{M\left(1-\frac{M}{K}\right)}{\sigma^2}\right) = \frac{30}{30-1}\left(1-\frac{21.34\left(1-\frac{21.34}{30}\right)}{(3.26)^2}\right) \approx 0.463$$

Where:

- KR-21: Reliability coefficient of the test
- K: Total number of questions on the test (30)
- M: Average score of all test-takers (21.34)
- σ^2: Variance of test scores (3.26^2)

A test is considered to have good reliability when the reliability coefficient is 0.7 or higher. The obtained result (KR-21 ≈ 0.463) suggests that the test lacks internal consistency among its questions or that the questions are not evenly distributed in terms of difficulty. The test requires revision, such as reviewing the content of the questions, improving the consistency among questions regarding learning objectives and assessment content, and removing overly easy or difficult questions.

3. Conclusion

The development of the MCQ bank for *Foundations of Mathematics in Elementary Education 1* followed a scientific and rigorous process, ensuring adherence to technical standards in test item construction. The pilot test of 30 multiple-choice questions yielded positive initial results but highlighted the need for further refinement, including adjusting difficulty levels, improving question content, and expanding testing to enhance validity, objectivity, and effectiveness. These improvements will not only strengthen assessment quality for this specific course but also contribute to broader educational enhancement, aligning with the ongoing demands of comprehensive educational reform. To ensure the MCQ bank's reliability, we continue to select statistically validated questions, revise those that do not meet standards, and conduct further testing, analysis, and evaluation of all items in the course. Once finalized, the MCQ bank will be officially implemented for student assessment under the credit-based system, providing a standardized and effective tool for evaluating learning outcomes in the coming years.

The findings of this study hold practical significance, particularly in their potential application to online learning environments such as Virtual Learning Environments. Assessment processes supported by virtual learning systems can help reduce the workload for instructors while maintaining fairness and objectivity in grading. Moreover, a well-constructed multiple-choice question bank can be instrumental in creating practice tests and mock exams, ultimately contributing to the enhancement of learning quality for students majoring in elementary education at HCMUE.

❖ ***Conflict of Interest:*** *Authors have no conflict of interest to declare.*

REFERENCES

Bailey, P. H., Mossey, S., Moroso, S., Cloutier, J. D., & Love, A. (2012). Implications of multiple-choice testing in nursing education. *Nurse Education today*, *32*(6), e40-e44. https://doi.org/10.1016/j.nedt.2011.09.011

Central Executive Committee. (2013). *Nghi quyet so 29-NQ /TW ngay 04/11/2013 ve Doi moi can ban, toan dien giao duc va dao tao, dap ung yeu cau cong nghiep haa, hien dai hoa trong dieu kien kinh te thi truong dinh huong xa hoi chu nghia va hoi nhap quoc te*. [Central Executive Committee. (2013). Resolution No. 29-NQ/TW dated November 4, 2013, on the Fundamental and Comprehensive Renovation of Education and Training to Meet the Requirements of Industrialization and Modernization in the Context of a Socialist-Oriented Market Economy and International Integration].

DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *Canadian Journal for the Scholarship of Teaching and Learning*, *2*(2), Article 4. https://doi.org/10.5206/cjsotl-rcacea.2011.2.4

Dang, A. H, Nguyen, T. T. A., Nguyen, D. T., & Nguyen. T. K. (2024). Thu nghiem cau hoi trac nghiem khach quan cho hoc phan giao duc hoc tai Truong Dai hoc Su pham Thanh pho Ho Chi Minh. [Multiple-choice question test bank for a course in Ho Chi Minh City University of Education]. *Ho Chi Minh City University of Education Journal of Science*, *21(1)*, 1-10. https://doi.org/10.54607/hcmue.js.21.1.4017(2024)

Dang, B. L. (2003). *Kiem tra đánh gia trong day - hoc dai hoc* [*Assessment in University Teaching and Learning*]. Education Publishing House.

Duong, T. T. (2005). *Trac nghiem va do luong thanh qua hoc tap (Phuong phap thuc hanh)* [*Multiple-choices and measurement of learning outcomes (Practical method)*]. Social Science Publishing House.

Faniyi, A. O. (2023). Enhancing student academic performance through educational testing and measurement. *Edumania-An International Multidisciplinary Journal*, *1*(2), 162-171. https://doi.org/10.59231/edumania/8981

Ho, V. L., Tran, T. H., Nguyen, D. D., & Nguyen, T. T. H. (2013). Quy trinh xay dung bo cau hoi trac nghiem kien thuc hoc phan giao duc hoc dai cuong theo chuong trinh dao tao tin chi o Truong Dai hoc Su pham Thanh pho Ho Chi Minh. [The process of building multiple-choices questions package of Education Foundation Course in credit curriculum at Ho Chi Minh Univerisity of Education]. *Ho Chi Minh City University of Education Journal of Science, 45,* 51-61. https://journal.hcmue.edu.vn/index.php/hcmuejos/article/view/1872

Ho Chi Minh City Universty of Education. (2024). *De cuong chi tiet hoc phan co so toan o tieu hoc 1* [*Fundamental mathematics for elementary teachers 1*].

Hoang, P. (2008). Tu dien tieng Viet [Vietnamese Dictionary]. Da Nang Publishing House.

Iñarrairaegui, M., Fernández-Ros, N., Lucena, F., Landecho, M. F., García, N., Quiroga, J., & Herrero, J. I. (2022). Evaluation of the quality of multiple-choice questions according to the students' academic level. *BMC Medical Education*, *22*(1), Article 779.

Jia, B., He, D., & Zhu, Z. (2020). *Quality and Feature of Multiple-Choice Questions in Education. Problems of Education in the 21st Century, 78 (4),* 576-594. https://doi.org/10.33225/pec/20.78.576

Lam, Q. T. (2008). *Trac nghiem va ung dung* [*Tests and Applications*]. Science and Technology Publishing House.

Liu, Q., Wald, N., Daskon, C., & Harland, T. (2023). Multiple-choice questions (MCQs) for higher-order cognition: Perspectives of university teachers. *Innovations in Education and Teaching International*, *61*(4), 802-814. https://doi.org/10.1080/14703297.2023.2222715

Ministry of Education and Training. (2021a). *Thong tu so 17/2021/TT-BGDĐT ngay 26/6/2021 Quy dinh ve chuan chuong trinh dao tao, xay dung, tham dinh va ban hanh chuong trinh dao tao các trinh do cua giao duc dai hoc* [*Circular No. 17/2021/TT-BGDĐT dated June 22, 2021, on regulations for program standards, development, appraisal, and issuance of training programs at higher education levels*]. https://vanban.chinhphu.vn/?pageid=27160&docid=203478

Ministry of Education and Training. (2021b). *Thong tu so 08/2021/TT-BGDDT Ban hanh Quy che dao tao trinh do dai hoc* [*Circular No. 08/2021/TT-BGDĐT Promulgating the Regulation on University Education*]. https://vanban.chinhphu.vn/default.aspx?pageid=27160&docid=203159

Tran, T. C. T., & Ngo, V. L. (2024). Mot so khuyen nghi xay dung de thi trac nghiem dap ung chuan dau ra hoc phan tren co so thang đo BLOOM. [Some recommendations for building multiple-choice tests that meet course outcome standards on the basis of BLOOM scale]. *Economy-Industry Journal, (39)*, 78-83. https://tcktcn.daihoclongan.edu.vn/index.php/ktcn/article/view/142

Tran, T. T. O. (2007). *Danh gia va do luong ket qua hoc tap* [*Assessment and Measurement of Learning Outcomes*]. University of Education Publishing House.

Vo, T. H. T., Tran, T. H., Nguyen, K. T., & Hoang, V. M. (2013). Xay dung bo cau hoi trac nghiem kien thuc hoc phan giao duc hoc pho thong theo chuong trinh dao tao tin chi o Truong Dai hoc Su pham Ho Chi Minh [Building multiple-choices questions package for General Education in credit curriculum Ho Chi Minh Univerisity of Education]. *Ho Chi Minh City University of Education Journal of Science, 48*, 114-123. https://journal.hcmue.edu.vn/index.php/hcmuejos/article/view/1915

XÂY DỰNG NGÂN HÀNG CÂU HỎI TRẮC NGHIỆM ĐỂ ĐÁNH GIÁ KẾT QUẢ HỌC TẬP HỌC PHẦN 'CƠ SỞ TOÁN Ở TIỂU HỌC 1' TẠI TRƯỜNG ĐẠI HỌC SƯ PHẠM THÀNH PHỐ HỒ CHÍ MINH

Nguyễn Việt Khoa**, *Phan Lê

Trường Đại học Sư phạm Thành phố Hồ Chí Minh, Việt Nam

**Tác giả liên hệ: Nguyễn Việt Khoa – Email: khoanvi@hcmue.edu.vn*

Ngày nhận bài: 25-01-2025; Ngày nhận bài sửa: 14-3-2025; Ngày duyệt đăng: 18-4-2025

TÓM TẮT

Đánh giá kết quả học tập của sinh viên là một yếu tố quan trọng trong giáo dục đại học, đóng vai trò then chốt trong việc đảm bảo và nâng cao chất lượng đào tạo. Nghiên cứu này tập trung vào việc xây dựng ngân hàng câu hỏi trắc nghiệm cho học phần Cơ sở Toán học ở Tiểu học 1 nhằm hỗ trợ đánh giá kết quả học tập của sinh viên. Nghiên cứu áp dụng phương pháp hỗn hợp (định tính và định lượng) dựa trên cơ sở lí thuyết về xây dựng đề kiểm tra và thiết kế câu hỏi trắc nghiệm. Quy trình xây dựng ngân hàng câu hỏi tuân theo năm bước khoa học chặt chẽ. Tổng cộng 300 câu hỏi trắc nghiệm đã được xây dựng ban đầu, trong đó 30 câu được thử nghiệm trên mẫu gồm 140 sinh viên. Kết quả thử nghiệm bước đầu xác nhận tính khả thi và hiệu quả ban đầu của ngân hàng câu hỏi, đồng thời chỉ ra một số điểm cần cải thiện. Cụ thể, theo chỉ số độ khó, có 13,3% câu hỏi được xếp loại khó, 23,3% mức trung bình và 63,3% mức dễ. Độ khó trung bình của bài kiểm tra là 0,7113—vượt nhẹ ngưỡng tối ưu (0,555–0,695)—cho thấy bài kiểm tra còn khá dễ đối với nhóm sinh viên mục tiêu. Những kết quả này cho thấy cần điều chỉnh lại sự phân bố mức độ khó của câu hỏi, đặc biệt là tăng tỷ lệ câu hỏi trung bình và khó, cải thiện độ rõ ràng và yêu cầu tư duy trong nội dung, đồng thời mở rộng phạm vi thử nghiệm trên các nhóm sinh viên đa dạng hơn. Những điều chỉnh này nhằm nâng cao khả năng phân hóa, độ tin cậy và sự phù hợp của bài kiểm tra với chuẩn đầu ra của học phần.

Từ khóa: đánh giá kết quả học tập; giáo dục toán học; ngân hàng câu hỏi trắc nghiệm; đánh giá khách quan