EVALUATING THE DIFFICULTY LEVEL OF OBJECTIVE MULTIPLE-CHOICE TEST QUESTIONS IN THE FOUNDATION MODULE IN THE PHYSICAL EDUCATION COURSE FOR STUDENTS AT THE VIETNAM NATIONAL UNIVERSITY OF AGRICULTURE

Nguyen Van Toan[1]; Le Trong Dong[1]
Nguyen The Hanh[1]

Summary

This study employs various research methods and identifies 54/100 usable questions, 46/100 questions that require modification or replacement. This assessment was based on the evaluation of the difficulty level of 10 multiple-choice theoretical test sets for the Foundation module in Physical Education. The research collected 460 theoretical test papers numbered from 1 to 10 in the mandatory Foundation module in the Physical Education course curriculum at the Vietnam National University of Agriculture (VNUA).

Keywords: Physical Education, multiple choice questions, difficulty level, Vietnam National University of Agriculture

INTRODUCTION

Students of the Vietnam National University of Agriculture (VNUA) are required to complete a Physical Education (PE) program consisting of 3 credits (1 mandatory credit and 2 elective credits). Each module consists of three grading components: study awareness (10%); mid-term exam (30%) in the form of multiple-choice questions; final exam (practical test) (60%).

The content of multiple-choice questions is divided into five types:

- Multiple-choice questions
- True-false questions (a special form of multiple choice)
- Matching questions
- Fill-in-the-blank questions
- Visual questions (a variation of fill-in-the-blank or matching questions using images)

MCQ, standing for Multiple Choice Question, is a widely used format in exams and surveys that presents several answer options. It is frequently used to obtain objective responses. An MCQ typically includes:

● A question: which must be precise, simple, and error-free

● A correct answer: the only accurate choice

● Distractors: other options designed to divert the respondent's attention

The theory of testing is an important branch of measurement science that is built on probability and statistics. It is applied in analyzing multiple-choice questions to improve their quality, eliminate or adjust flawed questions, and select questions based on set criteria.

In the process of implementing multiple-choice tests, constructing and evaluating the questions is crucial. The evaluation of test questions is carried out by determining their difficulty (Difficulty Factor) and discrimination (Discrimination Index) according to Robert L. Ebel's formula[*].

The mandatory Foundation module in the Physical Education course includes a mid-term exam that consists of a combination of multiple-choice formats. Therefore, the Department of Physical Education and the Center for Physical Education and Sports at VNUA need to assess the suitability of these objective test sets before using them to evaluate students.

The mandatory Foundation module in the Physical Education course is compulsory for

[1]Master, Vietnam National University of Agriculture

students at the Vietnam National University of Agriculture (VNUA). Each academic year, about 80 and 90 student groups participate in the course. The curriculum includes 30 sessions (2 periods in a session, 6 student self-prepared periods). Theoretical components are integrated into the sessions and are assessed through a mid-term exam (accounts for 30% of the total course grade) before a final examination (accounts for 60% of the total course grade). The mid-term exam is structured as an objective, multiple-choice test (MCQ), with a test set randomly selected by the Department of Physical Education following the VNUA's regulations and then distributed to the various student groups.

RESEARCH METHODS

The methods utilized in this research include:

- Document analysis and synthesis
- Pedagogical testing
- Statistical mathematics

The scientific basis for objective multiple-choice testing determines that question difficulty is classified as follows:

- Between 30% and 70% is "Acceptable" (with high discrimination ability);
- Between 15% and 30% or 70% and 85% indicates that the question needs revision;
- Below 15% or above 85% suggests the question is flawed and should be eliminated.

* The formula for calculating the Difficulty Factor (DF) [3]:

DF (%) = (H + L) × 100 / N. Where:

- H = Number of correct answers from the high-performing group
- L = Number of correct answers from the low-performing group
- N = Total number of students in both groups

The discrimination index of a question measures its effectiveness in distinguishing between students with high and low academic performance. The higher the discrimination index, the better the question performs in differentiating these groups. Generally, a question is considered positively discriminatory if higher-scoring students tend to answer it correctly while lower-scoring students tend to answer incorrectly.

A well-constructed multiple-choice test should include a significant proportion of questions with moderate difficulty. The average difficulty level for a question with n answer options is calculated as (100% + 1/n) / 2.

Generally, the difficulty of a test is "Acceptable": Average 0,25 - 0,75; "Too Easy" >0,75; "Too Difficult" <0,25

The ideal average score for a multiple-choice test is positioned between the maximum possible score (if all answers are correct) and the score that could be achieved through random guessing. For the mid-term PE exam at VNUA, the multiple-choice theoretical test sets are designed to consist of four-option multiple-choice questions. The DF is between 25% (¼ choose the correct answer randomly) and 100% (choose the correct answer on purpose), the ideal average score can be calculated as (10 + 2,5)/2,5 = 5.

If the observed average score deviates significantly from 5, the test can be classified as either "Too Easy" or "Too Difficult" for the students. In conclusion, if the ideal average score lies near the center of the observed score distribution, the test is considered to have an appropriate level of difficulty. Conversely, if the score is located toward either extreme of the distribution, the test may be classified as either "Too Difficult" or "Too Easy" for the student cohort.

RESULTS AND DISCUSSION

The results of the mid-term theoretical multiple-choice tests for the Foundation module in the Physical Education course at Vietnam National University of Agriculture (VNUA) during the first semester in the 2023-2024 academic year are presented here. A total of 460 students participated in the exams. The result of evaluating the difficulty index of each question in a test set is shown in Table 1.

Applying the difficulty index formula on test 1:

- "Too Easy": Question 9 with p=0,87 > 0,75. "Too Difficult": Question 1, and 5 with p<0,15, and Question 5 has p<0,25. These questions should be eliminated or changed the format.
- "Acceptable": Question 2, 3, 4, 6, 7, and 10 because the difficulty index is medium (0,25 - 0,75)

Table 1. Difficulty index of mid-term theoretical multiple-choice questions in the Foundation module in the Physical Education course (n=460), Test 1 - Test 10

Answer	Test 1		Test 2		Test 3		Test 4		Test 5	
Question	Correct answer	Difficulty index	Correct answer	Difficulty index	Correct answer	Difficulty index	Correct answer	Difficulty index	Correct answer	Difficulty index
1	A	0.07	C	0.22	C	0.23	A	0.35	D	0.07
2	B	0.37	D	0.26	D	0.85	B	0.75	D	0.67
3	D	0.26	D	0.09	D	0.96	D	0.1	D	0.13
4	C	0.5	B	0.85	B	0.81	C	0.4	A	0.47
5	A	0.11	C	0.07	C	0.38	C	0.4	C	0.44
6	C	0.39	A	0.11	A	0.04	D	0.73	D	0.51
7	A	0.5	D	0.37	D	0.94	D	0.9	C	0.8
8	D	0.37	C	0.46	C	0.96	B	0.85	B	0.24
9	C	0.87	C	0.07	C	0.81	A	0.29	B	0.44
10	A	0.43	A	0.02	A	0.35	A	0.42	B	0.2
Answer	Test 6		Test 7		Test 8		Test 9		Test 10	
1	A	0.31	B	0.2	C	0.09	D	0.14	C	0.44
2	D	0.86	C	0.62	C	0.21	D	1	D	0.62
3	B	0.48	D	0.18	B	0.58	C	0.31	D	0.02
4	C	0.83	B	0.36	B	0.35	C	0.46	B	0.13
5	C	0.21	B	0.22	D	0.53	D	0.6	C	0.78
6	B	0.62	D	0.64	D	0.88	D	0.91	A	0.33
7	B	0.69	C	0.44	D	0.81	D	0.77	D	0.13
8	A	0.33	C	0.62	C	0.98	B	0.26	C	0.4
9	B	0.45	C	0.51	C	0.33	C	0.66	C	0.56
10	A	0.48	D	0.33	A	0.49	C	0.49	A	0.47

Applying the difficulty index formula on test 2:

- "Too Easy": Question 4 with p=0,85 > 0,75. "Too Difficult": Question 1, 3, 5, 6, 9, 10 with p<0,25. These questions should be eliminated or changed the format. "Acceptable": Question 2, 7, and 8 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 3:

- "Too Easy": Question 2, 3, 4, 7, 8, 9 with p=0,85 > 0,75. "Too Difficult": Question 1, 6 with p<0,25. These questions should be eliminated or changed the format.

- "Acceptable": Question 5 and 10 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 4:

- "Too Easy": Question 7, 8 with p > 0,75. "Too Difficult": Question 3 with p<0,25. These questions should be eliminated or changed the format.

- "Acceptable": Question 1, 2, 4, 5, 7, 9, 10 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 5:

- "Too Easy": Question 7 with p=0,80 > 0,75. "Too Difficult": Questions 1, 3, 8, 10 with p<0,25. These questions should be eliminated or changed the format.

- "Acceptable": Question 2, 4, 5, 6, 9 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 6:

- "Too Easy": Question 2, 4 with p > 0,75. "Too Difficult": Question 5 with p<0,25. These questions should be eliminated or changed the format.

- "Acceptable": Question 2, and 4 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 7

- There is no "Too Easy" question in this test.

- "Too Difficult": Question 1, 3, 5 with p<0,25. These questions should be changed the format.

- "Acceptable": Question 2, 4, 6, 7, 8, 9, 10 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 8:

- "Too Easy": Question 6, 7 with p> 0,75. Question 8 with p=0,98. This question should be eliminated.

- "Too Difficult": Question 1= 0,09. This question should be eliminated. Question 2 with p=0,21<0,25. This question should be changed in the format.

- "Acceptable": Question 3, 4, 5, 9, 10 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 9:

- "Too Easy": Question 6 with p > 0,75. This question should be eliminated.

- "Too Difficult": Question 1 with p= 0,14. This question should be eliminated. Question 2, 8 with p<0,25. These questions should be changed the format.

- "Acceptable": Question 3, 4, 5, 7, 9, 10 because the difficulty index is medium (0,25 - 0,75)

Applying the difficulty index formula on test 10:

-n"Too Easy": Question 5 with p > 0,75. This question should be changed the format.

-n"Too Difficult": Question 3 with p= 0,02. This question should be eliminated. Question 4, 7 with p<0,25. These questions should be changed the format.

- "Acceptable": Question 1, 2, 6, 8, 9, 10 because the difficulty index is medium (0,25-0,75)

Conclusion: In objective multiple-choice tests, when conditions are consistent, scores tend to be more dispersed if many questions in the test fall within the moderate difficulty range. This suggests that many scores will be between scores that could be achieved through random guessing and the highest

possible score. In practice, multiple-choice questions, which were once considered effective, may no longer be suitable due to various factors. In terms of theory, once a multiple-choice question has been evaluated, standardized, and incorporated into a question bank, it generally does not need editing. Therefore, frequent assessment to determine the difficulty index and discrimination index of these questions is crucial in the deployment of objective tests in educational institutions in general, and specifically at the Vietnam National University of Agriculture.

In terms of general orientation, it's necessary for further research on the data and experimental verification to adjust questions, tests, and the difficulty and discrimination levels of each question and test. This will provide a solid scientific foundation to accurately assess students' competencies in the theoretical component of this module and the course as a whole.

CONCLUSION

The 1st 2023 - 2024 mid-term theoretical multiple-choice test for the Foundation module in the Physical Education course at Vietnam National University of Agriculture (VNUA) generally meets quality standards. Questions with a medium difficulty index (0,25-0,75) are considered usable. Specifically, in Test 1, the usable questions are Question 2, 3, 4, 6, 7, 10; in Test 2, the usable questions are Question 2, 7, 8; in Test 3, the usable questions are Question 5, 10; in Test 4, the usable questions are Question 1, 2, 4, 5, 7, 9, 10; in Test 5, the usable questions are Question 2, 4, 5, 6, 9; in Test 6, the usable questions are Question 1, 3, 6, 7, 8, 9, 10; in Test 7, the usable questions are Question 2, 4, 6, 7, 8, 9, 10; in Test 8, the usable questions are Question 3, 4, 5, 9, 10; in Test 9, the usable questions are Question 3, 4, 5, 7, 9, 10, in Test 10, the usable questions are Questios 1, 2, 6, 8, 9, 10

The remaining questions in all 10 tests include the "Too Difficult" questions $p<0,25$, and the "Too Easy" $p> 0.75$. These questions should be eliminated or changed to the questions, which have $p<0,15$.

The result of question difficulty and discriminative evaluation requires further comprehensive study to form a robust foundation for enhancing the quality of multiple-choice tests in the Foundation module in the Physical Education course at VNUA.

REFERENCES

1. Ministry of Education and Training (2015), Circular No. 25/2015/TT-BGDĐT dated October 14, 2015, by the Minister of Education and Training, Regulations on the curriculum for Physical Education courses in university-level programs, Hanoi.

2. Nguyen Van Canh, Nguyen Quoc Tuan (2021), Saigon University Science Journal, Evaluation of multiple-choice test quality through the application of the Rasch model and Conquest software, No. 75, pp. 52-63, March 2021.

3. Le Trung Thang (2022), Evaluation of multiple-choice questions through the determination of difficulty and discrimination indexes.

4. Lam Quang Thiep (2008), Tests and Applications, Science and Technology Publishing House.

Received 20/7/2024, Reviewed 24/9/2024, Accepted 28/11/2024
Main responsible: Nguyen Van Toan; Email: nvtoan@vnua.edu.vn)