

Test Item Analysis: A Catalyst for Quality Assurance of Assessment in Outcome-Based Education

Arati A. Pujari¹, Ketaki S. Sawashe², Snehal R. Watharkar³, Siddhanath V. Nishandar⁴.

^{1,2,3} Department of Electronics and Telecommunication Engineering, Rajarambapu Institute of Technology, Shivaji University, Sakharale, MS-415414, India.

⁴Department of Mechanical Engineering, Annasaheb Dange College of Engineering and Technology, Shivaji University, Ashta, MS-416301, India

¹arati.pujari@ritindia.edu ²ketaki.sawashe@ritindia.edu ³snehal.watharkar@ritindia.edu ⁴svn_mech@adcet.in

Abstract— Due to the evolving demands of the educational system, where a greater focus is placed on learning outcomes for students rather than merely degrees, outcome-based education (OBE) has grown in favor. An OBE system involves a clear picture of what is important for students to be able to do, organizing the curriculum, instruction and assessment to make sure this learning ultimately happens. Assessment and evolution is one of the important steps in Outcome based education. Purpose of the assessment and evolution is to final the lacuna and loopholes of teaching learning process and intended outcome that are finalized for particular course. Test item analysis is one of the prime components of assessment, which is the process of statistically analyzing student responses to individual exam questions. Facilitation value, discriminative index, and effectiveness of the distractor are the components of test item analysis. Item analysis helps in analyzing the difficulty level of items; whether the items are challenging, easy, comparatively difficult, or relatively easy, to discriminate amongst pupils who have different performance levels and identifies the effectiveness of the distractors in the multiple choice items.

Keywords—Item Analysis, Facilitation value, Discrimination Index, Effectiveness of the Distractors'

JEET Category—Practice

I. INTRODUCTION

The OBE idea places a focus on a curriculum with a pre-established set of learning outcomes. What kind of skill set students will have after graduating from an institution should be made apparent throughout the curriculum. According to Spady (1988), outcome-based education is a method of developing, delivering, and documenting instruction in terms of its intended goals and outcomes. OBE is further described by Spady (1994) as "Outcome-Based Education means clearly focusing and organizing everything in an educational system around what is essential for all students to be able to do successfully at the end of their learning experiences."

Fig.1 depicts the general flow of the teaching and learning process to follow OBE. Planning, delivery, and assessment are the components of this process. With the aid of the course syllabus and academic calendar, the first stage of preparation begins with course plan design. Next follows an assessment strategy, which may contain evaluation methods, duration, and outcomes to be attained through evaluation. Setting up test questions is also a part of it.

The second stage involves delivering course material utilizing various information and communication technology (ICT) tools, such smart boards, LCD projectors, Moodle, MS Teams, and others, while utilizing active learning resources, like Jigsaw, flipped classrooms, one minute papers, slido, kahoot, PBL, and many others. Additionally, it involves delivering knowledge by collaborating with industry through industry expert lectures, sponsored projects, and industry visits.

Last element is assessment which is nothing but evaluation of student's performance through in-semester evaluation and end semester evaluation of theory courses, practical courses as well as project work.

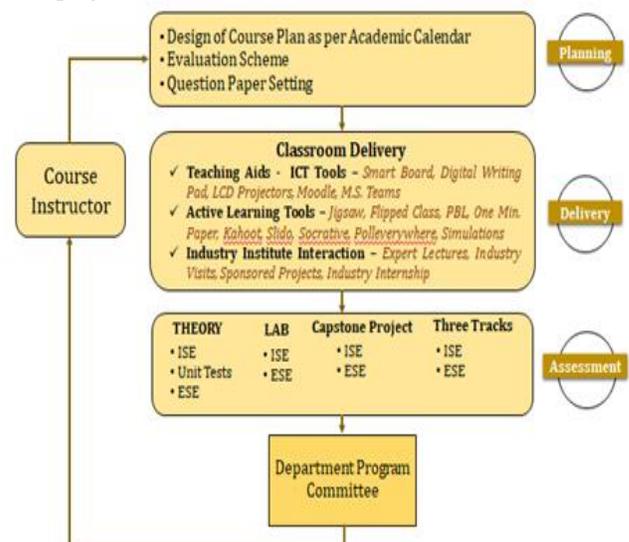


Fig.1. Process of Teaching-Learning

After assessment, attainment is computed and analyzed to determine whether the predetermined outcomes have been achieved. If the objectives are not met, the objectives or the way the content is delivered must be changed. So, the assessment plays a key part in all of these processes.

This paper was submitted for review on September, 10, 2023. It was accepted on November, 15, 2023.

Corresponding author: A. A. Pujari, Department of Electronics and Telecommunication Engineering, Rajarambapu Institute of Technology, Rajaramnagar, Islampur 415414, Maharashtra, India.

Address: Rajarambapu Institute of Technology, Rajaramnagar, Islampur 415414, Maharashtra, India. (e-mail: arati.pujari@ritindia.edu)

Copyright © 2024 JEET.

II. ASSESSMENT

One of the primary methods educators can learn more about students' learning is through assessments. Assessment is the process of gathering data. Once these data are gathered, teachers can then evaluate the student's performance. More specifically, assessment is the ways instructors gather data about their teaching and their students' learning. These assessments could be in the form of tests, quizzes, midterms, or final exams. Assessments also provide a key way for teachers to address students' needs, by showing education or performance gaps. Ideally, assessment data will help educators to figure out which topics students are struggling with and—more importantly—why. Assessments come in three types: diagnostic, formative, and summative. Despite the fact that all three are referred to as simply assessments, there are significant variations among them.

A. Diagnostic Evaluation

Before we begin teaching, a diagnostic examination will help us to determine our students present levels of understanding, their skill sets and capacities, and any misunderstandings they may have. We may make better plans for what to teach and how to teach it by being aware of the strengths and weaknesses of your students.

B. Formative assessment

Formative assessment is primarily used to gather information about students understanding and progress during the learning process. It is designed to provide ongoing feedback to both teachers and students to improve learning outcomes. Formative assessments are conducted throughout a course or instructional unit. They can be daily, weekly, or occur at various points during the learning process. Formative assessment includes Quizzes, short tests, assignments, class questions.

C. Summative Assessment:

Summative assessment is used to evaluate and summarize a student's overall learning and performance at the end of an instructional period, such as a course, semester, or academic year. Its primary purpose is to assign grades or determine whether learning objectives have been met. Summative assessment generally includes end semester examinations.

However, the job of the instructor is crucial. Pretend if we believed the students had attained the skills, but at the conclusion of the semester we discover that none of the pupils had. Since the time has passed, the students cannot receive a revision lesson at this moment. We can instead take periodic formative assessments, which may also serve as diagnostics, for each and every outcome during the teaching and learning process.

But again the effectiveness of an assessment depends not only on the content being tested but also on the quality of the questions or items used in the assessment. Test item analysis will help to analyze the quality of question, whether the questions are appropriate for the students or not. Item analysis refers to all the techniques used to assess the characteristics of test items and evaluate their quality during the process of test development and test construction

III. TEST ITEM ANALYSIS

The process of analyzing responses of particular test questions, or items, to determine whether or not their level of difficulty is acceptable is known as item analysis. This shows that the items are effective at discriminating amongst pupils who have different performance levels. Test item analysis is essential to maintaining the precision and fairness of tests. Although teachers frequently do it unintentionally, formalizing the procedure and outlining the steps involved offers a way to protect academic integrity and enhance assessments.

Test item analysis assists teachers in evaluating assessments to see whether they serve as legitimate means to test their students. For example, delivering an assessment that is either too challenging or too simple for a set of students is a waste of time and does not help us measure student learning. The frequent use of item analysis also allows teachers to evaluate examinations and identify any potential learning gaps. Teachers can then target and close those gaps by giving the appropriate teaching and support.

Components of item analysis: Three principal measures used in item analysis are facilitation value, discriminative index, and effectiveness of the distraction. Let's look at each of these factors and how they can help teachers to further understand test quality.

A. Facilitation value:

No test can exist without items, and each item will have a different level of difficulty. Facilitation value indicates the difficulty level of each test item i.e. question. Higher the number of students who are able to answer the question correctly, the easier the question (item) are. Therefore, the index of an item's difficulty is determined through calculating how many individuals could answer correctly from the number of test takers. If all of the students have provided the correct response, the facilitation value is 1, and if none of the students can, it is 0. As facilitation value ranges from 0 to 1, we need to compute intermediate values of facilitation value.

To obtain the intermediate facilitation value following steps need to be followed.

1. Rank the students based on the score secured.
2. Now distribute student into two groups; one is upper group and second is lower group.
3. Upper group consists of first 27% of the students i.e. 3dB down in the ranked list while lower group consists of last 27% of the students i.e. 3dB up in the ranked list. Students in upper group will have good performance than that of lower group.
4. Consider,

FV = Facilitation value

N = Size of a group

RU = Number of right answers given by upper group

RL = Number of right answers given by lower group

TABLE I
RESULT OF MULTIPLE CHOICE QUESTION TEST

Exam No.	Rank	Grade (15)	Question Number														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2005054	1	14	A	A	B	A	A	C	D	C	B	C	B	C	C	B	B
2005046	2	14	A	A	B	A	A	C	D	C	B	C	B	C	A	B	B
2005017	3	14	A	A	B	A	A	C	D	C	B	C	B	C	A	B	B
2005051	4	14	A	A	B	A	A	C	D	C	B	C	B	C	C	B	B
2005002	5	13	A	A	B	A	A	C	D	C	B	C	B	C	A	D	B
2005056	6	13	A	A	B	A	A	D	D	C	B	C	A	C	D	B	B
2005040	7	13	A	A	B	A	A	C	D	C	B	C	B	C	A	D	B
2005058	8	13	A	A	B	A	A	C	D	C	B	C	B	D	A	B	B
2006024	9	13	A	A	B	A	A	C	A	C	B	C	B	D	D	B	B
2005053	10	13	A	A	B	A	A	B	D	C	B	C	B	C	A	B	B
2005020	11	12	A	C	B	A	A	C	C	C	B	C	B	D	D	B	B
2005027	12	12	A	A	B	A	C	C	D	C	B	C	A	C	A	B	B
2005010	13	12	A	A	A	A	A	C	D	C	D	C	B	C	A	B	B
2005006	14	12	A	A	B	A	C	C	D	C	B	C	B	C	A	D	B
2155008	15	12	A	A	B	A	A	C	D	C	B	C	A	C	A	C	B
2005031	16	12	A	C	B	A	A	C	D	C	B	C	B	A	D	B	B
2005045	17	12	A	A	B	A	C	C	D	C	B	C	B	D	A	B	B
2005025	18	12	C	C	B	A	A	C	D	C	B	B	B	C	D	B	B
2005007	19	12	A	C	B	A	A	C	D	C	B	C	B	D	D	D	B
2005033	20	12	A	C	B	A	A	C	D	C	B	C	B	D	A	B	B
2005039	21	12	A	A	D	A	A	C	D	C	B	C	B	D	C	B	B
2005041	22	12	A	A	B	A	A	C	D	C	B	C	A	D	A	B	B
2155004	23	12	A	A	B	A	A	C	D	C	B	C	A	C	A	C	B
2005023	24	12	A	A	B	A	A	C	A	C	B	D	B	C	A	B	B
2005049	25	11	A	A	B	A	C	C	B	C	B	C	A	C	A	B	B
2005050	26	11	A	C	A	A	A	C	D	C	B	C	B	D	C	B	B
2005069	27	11	A	C	A	A	A	C	D	C	D	C	B	C	C	B	B
2005052	28	11	A	A	B	D	A	C	D	C	A	C	A	C	A	B	B
2005042	29	11	A	C	C	A	A	C	D	C	B	C	A	C	A	B	B
2005029	30	11	A	A	B	A	C	B	D	C	B	C	B	C	A	D	B
2005004	31	11	A	C	B	A	A	C	A	C	B	C	B	D	A	B	B
2005060	32	11	C	A	B	A	A	C	D	C	B	A	B	A	C	B	B
2005055	33	11	A	C	B	A	A	C	D	C	D	C	B	B	C	B	B
2005063	34	11	A	A	B	A	C	C	B	C	B	C	A	C	A	B	B
2005065	35	10	A	C	C	A	A	C	D	C	A	B	B	B	D	B	B
2005018	36	10	A	C	B	A	A	C	D	C	D	A	B	A	C	B	B
2005014	37	10	A	C	B	A	A	C	A	C	D	C	B	A	C	B	B
2155005	38	10	A	C	B	A	A	C	D	C	D	C	B	A	C	D	B
2005034	39	10	A	A	B	A	C	C	A	C	B	D	B	D	C	B	B
2009014	40	9	A	C	C	A	C	C	C	C	B	D	B	C	A	B	B
2155002	41	8	A	C	A	A	A	C	C	C	D	C	B	C	C	A	A
2005016	42	8	A	C	A	A	A	C	A	C	C	A	A	B	D	B	B
2005061	43	8	C	C	C	A	A	C	A	C	A	C	B	A	C	B	B
Correct Answer			A	A	B	A	A	C	D	C	B	C	B	C	D	B	B

Now by using following formula facilitation value (FV) can be calculated.

$$FV = \frac{\text{Total right answers recorded by students}}{\text{Total students of both group}}$$

$$\therefore FV = \frac{RU+RL}{2N} \quad (1)$$

According to Singh (2008) following factors may affect the facilitation value

1. The item may be challenging or ambiguous.
2. Previous knowledge or other experiences may introduce bias and lower the facilitation value.
3. It can also be lowered due to unfamiliarity about the type and content of items.
4. The type of options provided in case of multiple choice questions can also affect the facilitation value.

5. If every option is identical, it will be challenging for test-takers to choose the right one.

B. Discrimination Index:

The second component of test item analysis is the discrimination index. It is a measure of how well an item (i.e. a question) distinguishes between those with more skill (based on whatever is being measured by the test) and those with less skill. The discrimination index takes values between -1 and +1. Values close to +1 indicate that the item does a good job of discriminating between high and low performers. Values near zero indicate that the question does a poor job of discriminating between high and low performers. Finally, values near -1 indicate that the item tends to be answered correctly by those who perform the worst on the overall test and incorrectly by those who perform the best on the overall test.

By using the same notations mentioned in the previous session on calculating facilitation value, discrimination index (DI) can be obtained by using the following formula:

$$DI = \frac{RU-RI}{N} \quad (2)$$

Singh (2008) mentioned factors affecting the discrimination index as follow:

1. Since the discrimination index is closely tied to the facilitation value, the factors affecting the facilitation value likewise affect the discrimination index.
2. Heterogeneous test-takers.
3. The effectiveness of the distraction in affecting individuals who are unsure of the right response.

C. Effectiveness of the distractor

Multiple choice items significantly benefits for effectiveness of the distraction analysis. Distractors, as the name implies, are options that cause confusion for test-takers, particularly for those who don't know the right answer. Therefore, test takers must choose a distractor in order to understand its effectiveness; if unfortunately no test taker chooses the distractor, there is an urgent need to change or remove it. If DU= Number of student of the upper group chooses distractor, DL= Number of student of the lower group chooses distractor and N = Size of a group, then effectiveness of the distractor (ED) can be obtained as:

$$ED = \frac{DL-DU}{N} \quad (3)$$

IV. CASE STUDY

A multiple-choice question test was conducted for an elective course, Information Theory and Coding, in the third-year B. Tech. class. Table I gives the result of the same as per the rank of students. The question paper was designed with 15 multiple choice questions, each carrying one mark. Hence, the test taker can secure the highest mark of 15. Total 43 students appeared for the test; hence, the 3 dB group size is 12.

Test item analysis is performed on obtained result. Findings regarding components; facilitation value, discriminative index, and effectiveness of the distraction are mentioned in following sessions.

A. Facilitation value findings

Table II indicates the test items and there facilitation value. Taking into consideration that a facilitation value of 1 indicates that all students answered correctly and a facilitation value of 0 indicates that none of the students answered correctly, here are some conclusions made on test items:

Questions 4 and 8 have facilitation values of 1, indicating that all students answered them correctly. These questions appear to be very easy and may need to be made more challenging to effectively assess student knowledge.

TABLE II
TEST ITEMS AND THEIR FACILITATION VALUES

Question Number	1	2	3	4	5	6	7	8
Facilitation Value	0.92	0.58	0.79	1.00	0.83	0.92	0.63	1.00
Question Number	9	10	11	12	13	14	15	
Facilitation Value	0.67	0.75	0.83	0.50	0.21	0.83	0.96	

Question no. 13 have facilitation values 0.21, suggesting that very few students answered it correctly. This question may be too difficult or have issues that need to be addressed through revision.

Question no. 1, 6 and 15 has a high facilitation value of 0.92, 0.92 and 0.96 respectively, indicating that the majority of students answered them correctly. This suggests that the question may be relatively easy for the test-takers.

Questions 3, 7, 9, 10, and 14 have facilitation values ranging from 0.63 to 0.83, showing a mix of performance. These questions seem to provide a moderate level of challenge and effectively discriminate between students with varying levels of knowledge.

Questions 2 and 12 have facilitation values of 0.58 and 0.50, respectively, which are relatively lower. These questions might be challenging for students and may require review and potential revision.

B. Findings on discrimination index

Table II shows the test items and there discriminative indices.

TABLE III
TEST ITEMS AND THEIR DISCRIMINATION INDICES

Question Number	1	2	3	4	5	6	7	8
Discrimination index	0.17	0.67	0.42	0.00	0.17	-0.17	0.42	0.00
Question Number	9	10	11	12	13	14	15	
Discrimination index	0.67	0.50	0.00	0.50	0.08	0.00	0.08	

TABLE IV
TEST ITEMS AND EFFECTIVENESS OF

Question Number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Effectiveness of Distractor	A	✓	✓	0.17	✓	✓	0	0.25	0	0.17	0.25	0	0.42	-0.42	0.08	0.08
	B	0	0	✓	0	0	-0.08	0.08	0	✓	0.08	✓	0.25	0	✓	✓
	C	0.17	0.67	0.25	0	0.17	✓	0.08	✓	0.08	✓	0	✓	0.50	0	0
	D	0	0	0	0	0	-0.08	✓	0	0.42	0.17	0	-0.17	✓	-0.08	0

✓ Indicate correct answer of the question.

The item is effective in differentiating between high and low achievers when the values are close to +1. Values that are close to 0 indicate that it is difficult for the test item to distinguish between high and low achievers. Values close to -1 indicate that the item tends to be answered correctly by individuals who perform the worst on the overall test and incorrectly by those who achieve the best. According to these criteria we can draw the following conclusions:

Questions with Positive Discrimination (Values > 0):

Questions 2, 3, 7, 9, 10, and 12 have discrimination index values greater than 0. These values indicate that these questions do a good job of discriminating between high and low performers. Students who perform well on the overall test are more likely to answer these questions correctly, while students who perform poorly are less likely to answer them correctly.

Questions with Near-Zero Discrimination (Values \approx 0):

Questions 1, 5, 13 and 15 have discrimination index values that are close to zero (around 0.08 to 0.17). These values suggest that these questions do a poor job of discriminating between high and low performers. They are not effectively differentiating between students who perform well and those who perform poorly on the overall test.

Questions with exact Zero Discrimination (Values = 0):

Questions 4,8,11 and 14 have 0 discrimination index which indicate that these questions do a poor job of discriminating between high and low performers.

Questions with Negative Discrimination (Values < 0):

Question no. 6 has negative discrimination index. Negative discrimination value indicate that this test item tend to be answered correctly by those who perform poorly on the overall test and incorrectly by those who perform well. This suggests a significant issue with these questions, as they are not effectively assessing the intended knowledge or skills.

C. Findings on effectiveness of distractors

Distractors are the options in MCQ test that cause confusion for test-takers, particularly for those who don't know the right answer. Table IV shows test items 1, 2 and 3 with their effectiveness of distractors.

Following observations has implications for the effectiveness of these distractors:

Effectiveness of distractors B and D for test item 1 and 2 is zero. Similarly distractor D for test item 3 is zero. It suggests that these distractors are not serving their intended purpose.

Distractors are meant to confuse and challenge test-takers, particularly those who do not know the correct answer. When distractors are not chosen by any test-taker, it indicates a

potential issue with those response options. Educators should consider revising or removing these distractors to improve the quality of the assessment. As question number 4 and 8 has facilitation value of 1 indicating that all of the students answered correctly, effectiveness of all the distractors become zero.

V. CONCLUSION

For the multiple choice question tests item analysis concludes the importance of test items. Facilitation value plays important role in finding difficulty level of assessments. Questions that are too easy or too difficult may not effectively evaluate student knowledge, and adjustments may be necessary to align them with the intended learning outcomes. Discrimination index values provide insights into how well individual questions differentiate between high and low performers on the overall test. Questions with positive discrimination values effectively distinguish between students of varying abilities, while those with near-zero or negative discrimination values may need to be reviewed and potentially revised to improve their effectiveness in assessing student performance. Zero effectiveness of distractors suggests that distractors are not fulfilling their intended role in challenging test-takers. It is advisable to review and potentially revise or replace these distractors to improve the quality and effectiveness of the multiple-choice questions in the assessment. To determine whether students have attained particular learning outcomes, which is the main goal of outcome-based education, test item analysis aids in this process by offering insightful information on how well assessment tools, such as quizzes, measure the intended learning objectives. It offers educators practical insights they may apply to boost student learning, raise the standard of education overall, and refine teaching methods over time.

REFERENCES

- Obaydullah, A. K. M., Rahim, M. A., & Rahman, M. S. (2020). Concept of Outcome Based Education: Primary science. *Internation Journal of Advance Research And Innovative Ideas In Education*, 6(2), 1401-1407.
- Singh, A. K. (2015). Tests, measurements and research methods in behavioural sciences. Bharati Bhawan.
- Spady, W. G. (1988). Organizing for results: The basis of authentic restructuring and reform. *Educational leadership*, 46(2), 4-8.
- Spady, W. G. (1994). *Outcome-Based Education: Critical Issues and Answers*. American Association of School

Administrators, 1801 North Moore Street, Arlington,
VA 22209.

- Macayan, J. V. (2017). Implementing outcome-based education (OBE) framework: Implications for assessment of students' performance. *Educational Measurement and Evaluation Review*, 8(1), 1-10.
- Urbina, S. (2014). *Essentials of psychological testing*. John Wiley & Sons.
- Shyamal Kumar Das, "Outcome based Pedagogic Principles for Effective Teaching", NPTEL course.
- Lahrichi, A. (2019). Study on the effectiveness of formative and summative assessment techniques in education.
- Kumar, D., Jaipurkar, R., Shekhar, A., Sikri, G., & Srinivas, V. (2021). Item analysis of multiple choice questions: A quality assurance test for an assessment tool. *Medical Journal Armed Forces India*, 77, S85-S89.