International Journal of Informatics, Information System and Computer Engineering

Analysis of Question Items using the Differentiating Power Method

Fariz Zakaria*, Wasis Haryo Sasoko**, Ema Utami***

* Faculty of Informatics Engineering, Amikom University
** Faculty of Informatics Engineering, Amikom University
*** Faculty of Informatics Engineering, Amikom University
*Corresponding Email: farizakaria@students.amikom.ac.id

ABSTRACTS

The use of multiple choice questions in exams remains a common choice in education for various reasons, such as ease of assessment, perceived objectivity, and the ability to provide rapid feedback especially in large classes. Research shows that the use of multiple choice questions can strengthen retention of information, especially involving alternative wrong answers, as well as measure students' basic understanding in various subjects. However, to create quality multiple choice questions, an in-depth evaluation of the question elements is required, including item analysis to ensure the validity, reliability and fairness of the assessment. The results of this research using quantitative descriptive methods show that most of the questions can be improved, while a small number need to be rejected. The research conclusions suggest that rejected items should not be reused, while items that need to be corrected should be improved to improve the overall quality of the exam. Thus, analyzing the quality of multiple choice questions is crucial for increasing the effectiveness of assessment, especially in higher education contexts such as nursing and medical education.

ARTICLE INFO

1. INTRODUCTION

The use of multiple choice questions in exams is still a common choice for various reasons. From the teacher's perspective, multiple choice questions have advantages in terms of ease of assessment, perceived objectivity in assessment, can provide feedback in a short time, especially in large classes, and can also create more questions (Vanderbilt et al, 2013). In addition, research shows that the use of multiple-choice questions can help in the process of remembering information, especially related to alternative wrong answers. (Vanderbilt et al, 2013; Little et al, 2019) Multiple choice questions can also be used to measure students' basic understanding of material, such as in economics subjects (Stanger-Hall, 2012). Thus, the use of multiple choice questions in exams is still relevant because of the various advantages it offers, although it is also necessary to consider the weaknesses and implications of its use.

To create quality multiple choice questions, in-depth evaluation of the elements of the question is required. The importance of analyzing multiple choice questions includes several aspects such as providing educators with the ability to assess the extent to which questions can measure students' understanding and knowledge, especially in exams that have a significant impact on participants, such as in the field of Health (Stanger-Hall, 2012; Tarrant & Ware, 2012) [4]. Item analysis also helps identify weaknesses in question writing, test validity and reliability (Tarrant & Ware, 2008). Furthermore, good quality multiple choice questions can encourage deeper learning, strengthen student learning outcomes, and support the achievement of educational goals (Tarrant & Ware, 2008; Grainger et al, 2018). It is also important to carry out careful analysis to identify potential bias in the assessment of certain groups, such as male students (Grainger et al, 2018), to ensure equality in assessment. In addition, the analysis of multiple choice questions has a vital role in improving that the assessment can effectively measure core competencies in the educational context (Vlazneva & Androsova, 2021).

From the explanation above, we can conclude that, analysis of the quality of multiple choice questions is very important to ensure the validity, reliability, fairness, and effectiveness of assessments in various educational settings, especially in nursing, medical education, and other disciplines where assessments have high consequences. generally done.

2. METHOD

In this research, quantitative descriptive methods were used. The technique used for data collection is documentation. This method is used to retrieve research data in the form of a list of students working on questions, answer keys, and also a list of student answers.

The data analysis technique for the question items is differentiating power. Discriminating power is an assessment of the extent to which a question is able to show the difference between students in the upper group or students who are proficient and students in the lower group or students who are less proficient.

The following are the steps used in this research

2.1. Data Colection

The data used is a collection of exam data from various educational institutions obtained from the database of the application provider company. This data includes subjects, questions, answer keys, students taking exams, student answers

The following data is needed for this research

2.1.1. Data on subjects with the most test takers

> **select indexmapel, count(*) as number from cbt_dataanswer group by indexmapel order by number desc limit 10;**

This query will produce a table with two columns, namely indexmaple and total. Where the indexmapel is the subject and the number is the number of examinees. The table will be sorted from the highest number to the smallest.

From the data obtained, the data with the most data will be selected as the sample for this research, namely subjects with the index map: 78:1-G23-TRYOUT SNBT-02 with a total number of exam participants, namely 8893 students.

> **select * from answerdata where indexmapel="78:1-G23-TRYOUT SNBT-02" limit 10;**

2.1.2. Students answer data for the subject

The query returned all the answers from a total of 8893 students who took the exam

indexmapel	total
78:1-G23-TRYOUT SNBT-02	8893
557:557-SIMULASI-SAS-PAS-GANJIL	1057
148:SIMULASI_CBT_2023	966
231:TES OMRON	868
640:GANJIL 2023_XI BAHASA INDONESIA	829
640:GANJIL 2023_XI PPKN	826
640:GANJIL 2023_X BAHASA INGGRIS	823
640:GANJIL 2023_X BAHASA INDONESIA	822
640:GANJIL 2023 KELAS X SEJARAH NEW	820
640:GANJIL 2023 X PKN	819

Fig. 1. Subjects with the most test takers

indexkey	indexma pel	Answer
78:1-G23-TRYOUT SNBT-02:123	78:1-G23-TRYOUT SNBT-02	{"1":"D","2":"D","3":"D","4":"B","5":"C","6":"D","7":"D","8":"C","9":"E","10":"E","11":"E","12":"C","13":"B","14":"A","15":"E","16":"E","17":"B","18":"C","19":"D","20":"C","21":"C","22":"C","23":"A","24":"B","25":"B","26":"A","27":"C","28":"C","29":"B","30":"C","31":"C","32":"E","33":"D","34":"E","35":"D","36":"B","37":"E","38":"D","39":"B","40":"D","41":"D","42":"E","43":"E","44":"E","45":"E","46":"E","47":"E","48":"D","49":"B","50":"C","51":"C","52":"B","53":"B","54":"B","55":"A","56":"C","57":"A","58":"D","59":"C","60":"D","61":"B","62":"B","63":"B","64":"B","65":"A","66":"C","67":"D","68":"A","69":"A","70":"A","71":"D","72":"A","73":"E","74":"B","75":"A","76":"B","77":"C","78":"D","79":"D","80":"D","81":"C","82":"A","83":"C","84":"B","85":"A","86":"C","87":"B","88":"B","89":"E","90":"A","91":"A","92":"E","93":"C","94":"C","95":"E","96":"E","97":"B","98":"B","99":"B","100":"E","101":"C","102":"D","103":"B","104":"A","105":"B","106":"C","107":"D","108":"A","109":"C","110":"A","111":"A","112":"A","113":"B","114":"E","115":"D","116":"D","117":"B","118":"A","119":"E","120":"A","121":"A","122":"C","123":"D","124":"C","125":"D","126":"B","127":"B","128":"B","129":"C","130":"D","131":"D","132":"A","133":"C","134":"E","135":"C","136":"C","137":"D","138":"A","139":"B","140":"D","141":"A","142":"B","143":"D","144":"C","145":"C","146":"B","147":"B","148":"D","149":"B","150":"E","151":"C","152":"B","153":"A","154":"A","155":"B"}
78:1-G23-TRYOUT SNBT-02:G0052426134	78:1-G23-TRYOUT SNBT-02	{"1":"E","2":"A","3":"E","4":"D","5":"D","6":"B","7":"D","8":"C","9":"D","10":"D","11":"E","12":"C","13":"A","14":"A","15":"C","16":"A","17":"B","18":"C","19":"D","20":"E","21":"A","22":"C","23":"B","24":"B","25":"B","26":"A","27":"C","28":"A","29":"B","30":"B","31":"A","32":"A","33":"D","34":"E","35":"D","36":"B","37":"C","38":"E","39":"C","40":"A","41":"E","42":"D","43":"D","44":"A","45":"E","46":"A","47":"D","48":"C","49":"A","50":"D","51":"B","52":"D","53":"E","54":"D","55":"E","56":"D","57":"A","58":"D","59":"C","60":"E","61":"D","62":"A","63":"B","64":"A","65":"A","66":"C","67":"D","68":"A","69":"A","70":"A","71":"B","72":"B","73":"A","74":"A","75":"E","76":"D","77":"A","78":"A","79":"C","80":"D","81":"D","82":"B","83":"C","84":"A","85":"A","86":"A","87":"B","88":"A","89":"B","90":"C","91":"A","92":"B","93":"C","94":"A","95":"B","96":"A","97":"D","98":"C","99":"B","100":"E","101":"A","102":"A","103":"E","104":"A","105":"B","106":"A","107":"C","108":"A","109":"E","110":"C","111":"A","112":"A","113":"B","114":"B","115":"A","116":"A","117":"A","118":"B","119":"D","120":"A","121":"B","122":"B","123":"A","124":"D","125":"B","126":"E","127":"B","128":"C","129":"C","130":"B","131":"B","132":"B","133":"C","134":"C","135":"A","136":"C","137":"D","138":"A","139":"B","140":"C","141":"C","142":"D","143":"C","144":"E","145":"C","146":"C","147":"C","148":"C","149":"C","150":"C","151":"E","152":"C","153":"A","154":"B","155":"A","156":"A","156001":"0895600456246"}

Fig 2. Students answer data for the subject

2.1.3. Data student who work

```
select * from datakey where
indexkey = "78:1-G23-TRYOUT
SNBT-02"
```

This query produces the answer key for the sample questions

indexkey	key
78:1-G23-TRYOUT SNBT-02	{"kunci": {"1":"C","2":"D","3":"A","4":"B","5":"C","6":"B","7":"A","8":"B","9":"C","10":"C","11":"B","12":"B","13":"B","14":"A","15":"D","16":"A","17":"B","18":"C","19":"D","20":"C","21":"C","22":"C","23":"B","24":"B","25":"B","26":"A","27":"C","28":"A","29":"B","30":"A","31":"C","32":"D","33":"E","34":"E","35":"C","36":"B","37":"E","38":"A","39":"B","40":"B","41":"A","42":"E","43":"C","44":"D","45":"D","46":"E","47":"E","48":"D","49":"E","50":"A","51":"B","52":"C","53":"A","54":"A","55":"E","56":"B","57":"D","58":"E","59":"C","60":"D","61":"A","62":"B","63":"B","64":"B","65":"A","66":"C","67":"D","68":"B","69":"A","70":"A","71":"A","72":"C","73":"D","74":"A","75":"A","76":"B","77":"D","78":"C","79":"B","80":"A","81":"B","82":"C","83":"C","84":"B","85":"C","86":"C","87":"A","88":"B","89":"D","90":"B","91":"A","92":"B","93":"C","94":"C","95":"B","96":"A","97":"B","98":"B","99":"B","100":"C","101":"B","102":"A","103":"E","104":"A","105":"B","106":"A","107":"C","108":"B","109":"E","110":"C","111":"A","112":"C","113":"B","114":"D","115":"A","116":"E","117":"C","118":"B","119":"D","120":"C","121":"E","122":"B","123":"E","124":"A","125":"B","126":"E","127":"E","128":"B","129":"C","130":"B","131":"E","132":"B","133":"D","134":"A","135":"D","136":"E","137":"A","138":"B","139":"C","140":"A","141":"D","142":"B","143":"A","144":"A","145":"D","146":"D","147":"B","148":"B","149":"C","150":"D","151":"C","152":"D","153":"A","154":"A","155":"B","156":"A","156001":"X"},"bobot": {"1":"1","2":"1","3":"1","4":"1","5":"1","6":"1","7":"1","8":"1","9":"1","10":"1","11":"1","12":"1","13":"1","14":"1","15":"1","16":"1","17":"1","18":"1","19":"1","20":"1","21":"1","22":"1","23":"1","24":"1","25":"1","26":"1","27":"1","28":"1","29":"1","30":"1","31":"1","32":"1","33":"1","34":"1","35":"1","36":"1","37":"1","38":"1","39":"1","40":"1","41":"1","42":"1","43":"1","44":"1","45":"1","46":"1","47":"1","48":"1","49":"1","50":"1","51":"1","52":"1","53":"1","54":"1","55":"1","56":"1","57":"1","58":"1","59":"1","60":"1","61":"1","62":"1","63":"1","64":"1","65":"1","66":"1","67":"1","68":"1","69":"1","70":"1","71":"1","72":"1","73":"1","74":"1","75":"1","76":"1","77":"1","78":"1","79":"1","80":"1","81":"1","82":"1","83":"1","84":"1","85":"1","86":"1","87":"1","88":"1","89":"1","90":"1","91":"1","92":"1","93":"1","94":"1","95":"1","96":"1","97":"1","98":"1","99":"1","100":"1","101":"1","102":"1","103":"1","104":"1","105":"1","106":"1","107":"1","108":"1","109":"1","110":"1","111":"1","112":"1","113":"1","114":"1","115":"1","116":"1","117":"1","118":"1","119":"1","120":"1","121":"1","122":"1","123":"1","124":"1","125":"1","126":"1","127":"1","128":"1","129":"1","130":"1","131":"1","132":"1","133":"1","134":"1","135":"1","136":"1","137":"1","138":"1","139":"1","140":"1","141":"1","142":"1","143":"1","144":"1","145":"1","146":"1","147":"1","148":"1","149":"1","150":"1","151":"1","152":"1","153":"1","154":"1","155":"1","156":"1","156001":"1"},"locking":{},"kd":{},"kdsoal":{},"grouping":{}}

Fig. 3. Answer key for the sample questions

2.2. Data Processing

The data obtained is semi-structured data. To analyze this data, the data is converted into structured data.

2.3. Data Analysis

After the data is converted into structured data, the data is then analyzed using the Differentiating Power Technique.

3. RESULTS AND DISCUSSION

After all the data has become structured data, the data is then analyzed using the Python programming language to determine the differentiating power of each item in the research sample. Before calculating the Differentiating Power, do the following steps:

3.1. Student Data is sorted based on the scores obtained

Fig. 4. Student data scores.

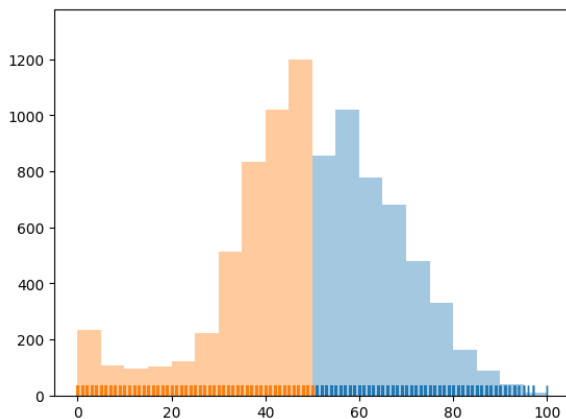### 3.2. Sorted data. Then Devided into the upper group and lower group

Fig. 5. Upper and Lower Group Data

3.3. Analysis of Differentiating Power of Question Items

To analyze the differentiating power of the test items, a formula is used

$$DP = \frac{2(K_a - K_b)}{n}$$

DP = Differentiating Power
Ka = Number of students in the top group who answered correctly
Kb = Number of students in the lower group who answered correctly
n = Number of students

Table 1. Differentiating Criteria.

Differentiating Criteria	Description
DP > 0.25	Accepted
0 < DP ≤ 0.25	Corrected
DP ≤ 0	Rejected

After carrying out the calculations, the following results were obtained

Fig. 6. Result of Analysis

Graph of the results of calculating the Discriminating Power of Question Items
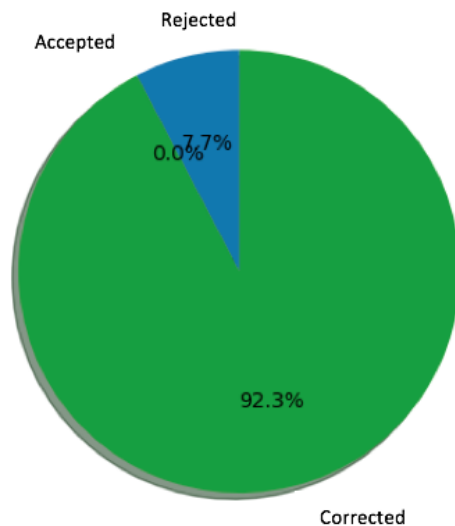
Fig. 7. Graph of the Results.

Based on the results of calculating the differentiating power, 14 questions or 7.7% of the criteria were rejected, 144 questions or 92.3% of the criteria were corrected and there were no questions or 0% of the criteria accepted.

Criteria	Index DP	Question Item	Total	Percentage
Accepted	DP > 0.25		0	0%
Corrected	0 < DP £ 0.25	[0, 1, 2, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 40, 41, 43, 44 , 45, 46, 48, 49, 50, 51, 52, 53, 54, 56,57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 68, 69, 70, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 150, 151, 152, 153, 154, 155]	144	92.3%
Rejected	DP £ 0	[3, 6, 32, 39, 42, 47, 55, 67, 71, 92, 130, 149]	12	7.7%

Fig 9. Results of calculating the differentiating power of question items

4. CONCLUSION

Based on the research results, it was concluded that question 78:1-G23-TRYOUT SNBT-02 was of poor quality because there were 144 questions or 92.3% of which had revised criteria and there were 12 questions or 7.7% of which had rejected criteria. , no questions were even accepted.

Question items with rejected criteria should not be used again for future tests, while questions with revised criteria should be improved to make them better questions.

ACKNOWLEDGMENTS

REFERENCES

[1] Vanderbilt, A., Feldman, M., & Wood, I. (2013). Assessment in undergraduate medical education: a review of course exams. *Medical education online*, *18*(1), 20438.

[2] Little, J. L., Frickey, E. A., & Fung, A. K. (2019). The role of retrieval in answering multiple-choice questions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(8), 1473.

[3] Stanger-Hall, K. F. (2012). Multiple-choice exams: an obstacle for higher-level thinking in introductory science classes. *CBE—Life Sciences Education*, *11*(3), 294-306.

[4] Tarrant, M., & Ware, J. (2012). A framework for improving the quality of multiple-choice assessments. *Nurse Educator*, *37*(3), 98-104.

[5] Tarrant, M., & Ware, J. (2008). Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical education*, *42*(2), 198-206.

[6] Grainger, R., Dai, W., Osborne, E., & Kenwright, D. (2018). Medical students create multiple-choice questions for learning in pathology education: a pilot study. *BMC medical education*, *18*, 1-8.

[7] Vlazneva, S., & Androsova, O. (2021). Multiple-choice questions and essays in assessing economics. In *SHS Web of Conferences* (Vol. 99, p. 01032). EDP Sciences.