International Journal of Multicultural and Multireligious Understanding

http://ijmmu.com
editor@ijmmu.com
ISSN 2364-5369
Volume 11, Issue 10
October, 2024
Pages: 58-71

Item Analysis of School Examination Questions Using the Iteman Program

Ghina Kamilah Siregar; Prihadi; Burhan Nurgiyantoro

Magister of Education Indonesian Language and Literature, Faculty of Language, Arts, and Culture, Yogyakarta State University, Indonesia

http://dx.doi.org/10.18415/ijmmu.v11i10.6153

Abstract

This study aims to analyze and describe question items based on validity, reliability, difficulty index, differentiation index and the effectiveness of distractors on the final exam of the odd semester of class X SMA Kartika 1-5 Padang in the 2023/2024 school year. The method used is descriptive quantitative. Data obtained in the form of answer sheets of 23 students of SMA Kartika 1-5 Padang using the Iteman program. The results showed that the reliability of the questions carried out on students obtained an Alpa Cronbach score of 0.830 in the Iteman program. The question is declared reliable because the reliability index is above 0.60. Based on the Differentiation Index, it is known that 23 questions are categorized as feasible, 7 questions are categorized as inappropriate and need to be revised. Then based on the Difficulty Level Index, it is known that 5 questions are at a difficult level, 21 questions at an easy level and 5 questions at a very easy level.

Keywords: *Item Analysis; Iteman Program; School Exam Questions*

Introduction

A teacher must be able to plan assessment or evaluation actions, conduct evaluations, and report evaluation or assessment results (Rasna et al., 2023).So that learning assessment will provide information on the success of student learning. This is in line with the opinion (Nurgiyantoro, 2017) which states that assessment is an educational and teaching activity which is a process in achieving a number of predetermined goals.

An assessment is needed to see the achievement of students' competence in the learning process, this is a form of implementation of Indonesian language policy (Suharjo & Sumardi, 2024). (Sarker & Ullah, 2023) stated that the quality of education can be seen through the process dimension which includes learning, subject content, teacher knowledge, accuracy of curriculum content, learning media, and student assessment. Therefore, assessment must be carried out and planned properly. The party that plays a role and is responsible for giving assessments in this case is the teacher or educator.

As a party who gives an assessment, educators must understand the assessment itself. Assessment requires data obtained from measurement, while measurement requires assessment as the purpose of measurement. (Nurgiyantoro, 2017) suggests that measurement is a process to obtain a description in the

form of numbers or scores that show the level of achievement of a person or student in a field. Meanwhile, assessment is a systematic process of collecting, analyzing, and interpreting information to determine how far a student can achieve educational goals.

Assessment is not a simple matter of just assessing learner outcomes. There are many activities in assessment activities and not only assessing student work but also the quality of learning. The assessment instrument used must be considered. Before assessing, achievement indicators are formulated first by describing basic competencies. One of the assessment instruments that can be used by educators is a test (Dachliyani, 2019). The results of the assessment are then used as reference material for evaluation. (López-Hernández et al., 2023) see the need to link assessment and feedback to student learning outcomes, rather than using it only as a tool to check standards at the end of learning.

As one of the tools for measuring learning outcomes, tests are expected to be able to provide information that can be accounted for. Test is one of the assessment tools in written form to record or observe student achievement in line with the assessment target. When making evaluation tools we are asked to determine the level of items written into easy, medium, and difficult categories. Determining the difficulty of the items is not only based on approximate considerations, but needs to be calculated through empirical data of a test tool being tested. Based on these results, the index of the level of difficulty and the effectiveness of the test question distractors that have been made can be determined. A good question is a question whose difficulty level is not too easy and not too difficult (Amalia & Widayati, 2012). The level of difficulty of the items also has an influence on the differentiation of the questions.

In order to achieve goals in the learning process, it is necessary to know the quality of the test tools that are done by students. The quality of questions or test tools can be determined through item analysis. The purpose of item analysis according to (Arikunto, 2018) is to find good, good enough, and poorquality items and to get information to improve or perfect related questions.

In addition, there is a previous study entitled Analysis of school exam items using the Iteman program conducted by (Pradani & Efendi, 2023), this study shows the category of items based on content validity, reliability, difficulty level, question differentiation and the operation of distractors on junior high school exam sheets in Rembang Regency. Then further research was conducted by (Amalia & Widayati, 2012) entitled Analysis of Quality Control Test Items for Class XII High School Accounting Economics Subjects in Yogkayarta City in 2012, the study described item validity, reliability, difficulty level and effectiveness of distractors on quality control test questions for class XII high school. These two studies are interesting and relevant to be used as references because they both examine item analysis which includes validity, reliability, difficulty level and the effectiveness of distractors on test questions.

The focus of this research is to analyze the quality of the end-of-semester items of Indonesian language class X using the Iteman program. The author wants to examine the quality of Indonesian language items in the final semester exam at SMA Kartika 1-5 Padang. This test serves to measure the level of development or progress that students have achieved after they have gone through the learning process. The question items analyzed were tested on grade XI students at SMA Kartika 1-5 Padang as an assessment of the final semester exam in the 2023/2024 school year. The questions tested were done by 23 students. The test used google form and was conducted offline at school.

Literature Review

The establishment of Indonesian as a subject in every educational unit is one of the policies in the world of education in Indonesia (Asri, 2017). Indonesian language learning refers to the Indonesian language proficiency standards contained in the content standards and competency standards for graduates of each level of education. For the realization of educational goals, language learning and assessment policies are important (Ilyas, 2016). This is also in line with (Danny, 2012) which states that

the quality of education can be seen from the process dimension, namely learning, subject content, teacher knowledge, accuracy of curriculum content, learning media and student assessment.

The question items that are prepared must meet the requirements and writing rules. The requirements for the preparation of multiple choice questions in the Ministry of Education must pay attention to three aspects consisting of material, construction, and language aspects (Nurgiyantoro, 2017). First, in the material aspect, it is important to pay attention to the suitability of the question with the indicators to be achieved, meaning that the question must ask or consider the behavior and material to be measured based on the formulation of indicators in the grid. In addition, the exception options must work, and each question must have one correct answer, meaning that in one question only has one answer key. Second, the construction aspect, namely the subject matter must be formulated clearly and clearly. This means that the material to be measured in the question must be clear and not cause different interpretations. Next, each item that is prepared only contains one problem. Third, in the language aspect, each item must use language that is in accordance with the rules of Indonesian language writing.

The question items prepared for assessment also need to be considered for quality. Things that need to be considered when analyzing items are calculating the level of difficulty of the items, the differentiation of the items, and the effectiveness of the items (distractors). The level of difficulty of a question item can be said with an index ranging from 0.00-1.00. An index of 0.00 means that the item is in the very difficult category because no one student can answer it correctly. Conversely, an index of 1.00 means that the item is in the very easy category because all students are able to answer correctly. (Daryanto, 2013) states that all items can be called feasible if the difficulty index ranges from 0.15-0.85.

Method

This research uses a quantitative descriptive method. Quantitative descriptive in this study is used to describe the results of obtaining scores on the Iteman program. The data sources of this research are grids, question sheets, answer keys, and student answer sheets. Furthermore, the analysis was carried out on these documents to be reviewed for content validity, reliability, difficulty level, differentiation index, and the effectiveness of the item distractors using the Iteman program.

The analysis was carried out on 30 items of the final semester exam in Indonesian language subject class X SMA Kartika 1-5 Padang in the 2023/2024 school year. The stages of the research carried out were first, the researcher gave a questionnaire to the teacher who taught Indonesian language subjects in class X SMA Kartika 1-5 Padang to review the suitability between the question grids and the question sheets that had been prepared by the researcher. Then improvements were made after the questions were reviewed and the questions were distributed to students to do. Later, the answer sheet is used as material for analysis.

After this process, the question reliability analysis was carried out which was obtained from the Iteman program analysis output in the Alpha Cronbach section. Iteman provides statistical data obtained from the data that has been entered. The Index of Difficulty Level (ITK) was obtained through the proportional correct section as a result of Iteman calculation. The Index of Distinctiveness (IDB) is known through the point biserial as the result of the Iteman program calculation. The effectiveness of the items can be seen in the proportional endorsing score in the Iteman program.

After the research data in the form of student answers and answer keys were entered in excel, data analysis was carried out using the Iteman program. In the Iteman program in the Windows Files section, select the data matrix file and then fill it in with the student answer data which has been changed to Text (Tab delimited) format. In the Item control file section, fill in the answer key of the related question, and in the output file section, write the file name that will appear in excel and word. Furthermore, in the run title section, the file name is also written.

Next, in the Input Format (Fixed Width Data) windows, the Number of examinee ID columns section is filled with the number of IDs, namely 23. Then in the Examinee IDs begin in column section, the number 1 is filled in (meaning it starts from the 1st column). Furthermore, the Item responses begin in column section is filled with the number 7 (meaning the response starts from the 7th column). Then move to the Delimited Data section by clicking the checkbox in the The data matrix file is delimited by a, then click Tab. Then move to the Scoring Options window and click on the Total Score, Each domain separately, and Standardized, with a mean of which will automatically appear a check mark. Then the last stage, in the Output Options section click Run then select yes and the data is finished analyzing and automatically saved in a separate folder.

Result and Discussion

Validity of Question Grids and Question Items

Before the questions are tested on students, the grids and questions are reviewed to ensure that the items written have met the demands of a good question, the process of reviewing grids and items is a step that reflects how valid the items to be tested (Amalia & Widayati, 2012). Reviewing these items allows for revisions so that the questions tested to students are the best questions and are truly able to measure the abilities of students. The validity used in this analysis is content validity, which is the validity that must be met in the test tool and has the essence of determining the accuracy of sampling the teaching material to be passed on (Nurgiyantoro, 2017). Validity was carried out by a peer, namely Fioca Rifka Fortuna, S.Pd. As the Indonesian language subject teacher at SMA Kartika 1-5 Padang.

Table 1. Validity Results of Question Numbers 1-30

Aspects	**Statement Type**	**Item Number**		
		1-10	**11-20**	**21-30**
Content	Question items are in accordance with the indicators	√	√	√
	The content is scientifically correct	√	√	√
	Only one correct answer key	√	√	√
	Content is appropriate for the grade/level of education	√	√	√
	Exception items are functioning properly	√	√	√
Construction	Subject matter is clearly formulated	√	√	√
	The subject matter does not lead to the correct answer	√	√	√
	Answer options are clearly formulated	√	√	√
	Homogeneous answer options	√	√	√
	No double negative form	√	√	√
	The length of the answer options is approximately the same	√	√	√
	Items are not dependent on each other	√	√	√
	Options in number/time sorted	√	√	√
Language	Communicative language	√	√	√
	Grammatical sentences	√	√	√
	Sentences are not double-meaning	√	√	√
	Standard/general/neutral vocabulary	√	√	√

Based on the review sheet conducted by peers above, it can be concluded that the 30 multiple choice questions tested are valid questions. This can be seen from several aspects, namely the material, construction and language aspects of each item prepared. Content validity refers to an understanding of the test instrument in accordance with the objectives and description of the subject matter that has been studied. A test can be said to be valid if the questions are clearly arranged to measure certain learning competency indicators that represent all teaching materials that have been studied. Therefore, it can be concluded that the final semester exam items tested on 23 students of class X SMA Kartika 1-5 Padang in the 2023/2024 academic year have a high level of validity.

Reliability Index

Reliability refers to the notion of measurement consistency, namely how consistent the cortus or evaluation results are from one measurement to another. High and low reliability will affect validity. High reliability will allow the achievement of validity. Reliability refers to the results obtained that are tested with a test instrument and not the test tool itself. The reliability used in this study is the Alpha Cronnbach reliability of the Iteman program.

Item analysis carried out using Iteman will provide statistical data (scale statistics). The results of this reliability indicate an index of the extent to which a measuring instrument can be trusted or reliable. From this statistical data, information will be found about the magnitude of the reliability coefficient calculated by Cronbach's Alpha, the average score, minimum score, maximum score, and average ITK (Mean P) for all items, and average IDB (Mean R) for all items. The analysis for this assessment uses Iteman. The statistical data obtained from Iteman are as follows.

Table 2. Results of Reliability Analysis Using Iteman

Items	30
Mean	21.,739
SD	5,285
Min Score	11
Max Score	29
Mean P	0,725
Mean Rpbis	0,330
Reliability Analysis Statistic, Scored Items Alpha: 0,820	

Based on statistical data with 30 items tested on students, it can be obtained that the Cronbach Alpha coefficient is 0, 820. The average difficulty index (ITK) of all items is 0.725. The average index of distinguishing power (IDB) is 0.330. A test can be said to be reliable if it has a coefficient of more than 0.60. Based on these assumptions, the test tool tested in this assessment is declared reliable because the amount of the reliability index is more than 0.60, namely 0.820.

Index of Difficulty (ITK) and Index of Distinction (IDB)

As previously described, the item analysis used in this assessment uses Iteman. From 30 items, ITK and IDB data were obtained. The index of difficulty (ITK) is an index that shows how easy or difficult an item is for the test taker being tested. A good item is one whose difficulty level is sufficient, not too easy or too difficult. The tolerated ITK is between 0.20 - 0.80. The ITK scale classification can be seen in the following table.

Table 3. Index of Difficulty Classification (ITK)

Index of Difficulty	
Criteria	**Scale**
0,00 – 0,20	Very difficult
0,21 – 0,40	Difficult
0,41 – 0,60	Medium
0,61 – 0,81	Easy
0,81 – 1,00	Very easy

Item differentiation is a statement about how much power an item can distinguish between high and low group participants. The index of distinguishing power (IDB) is an index that shows how much power an item has between high and low group participants. Theoretically, the IDB can range from -1.00 to +1.00. However, indices that detect 0 (zero) or negative numbers are declared inappropriate. A question item is said to be feasible if it has an IDB of at least 0.20. The IDB classification can be seen in the following table.

Table 4. Index of Distinction Classification (IDB)

Index of Distinction	
Criteria	**Scale**
0,40 – 1,00	Worth
0,30 – 0,39	Medium (feasible and no need to revise)
0,20 – 0,29	Medium (feasible and needs to be revised)
Negative – 0,19	Bad (not feasible)

The results of the ITK and IDB calculations in this assessment were not carried out manually, but were carried out with the Iteman program. The results of the level of difficulty and feasibility of each item can be seen in the following summary table.

Table 5. Results of Item Analysis of Final Semester Question of Indonesian Language Class XI SMA Kartika 1-5 Padang, Year 2023/2024 Based on ITK and ITB

Question Number	ITK	IDB	Description
1	0,696	0,139	Easy and Bad (Not feasible)
2	0,522	0,612	Medium and Feasible
3	0,826	0,318	Very easy and feasible (no revision needed)
4	0,609	0,220	Sedang dan Layak (perlu direvisi)
5	0,826	0,318	Very easy and feasible (no revision needed)
6	0,826	0,084	Very easy and Bad (Not feasible)
7	0,783	0,473	Easy and Feasible
8	0,783	0,473	Easy and Feasible
9	0,826	0,153	Very easy and Bad (Not feasible)

10	0,652	0,311	Easy and Medium (no need to revise)
11	0,783	0,450	Easy and Feasible
12	0,696	-0,233	Easy and Bad (Not feasible)
13	0,783	-0,170	Easy and Bad (Not feasible)
14	0,739	0,323	Easy and Medium (no need to revise)
15	0,696	-0,161	Easy and Bad (Not feasible)
16	0,739	0,261	Easy and Medium (needs revision)
17	0,696	0,641	Easy and Feasible
18	0,696	0,374	Easy and Medium (no need to revise)
19	0,696	0,043	Easy and Bad (Not feasible)
20	0,522	0,612	Medium and Feasible
21	0,696	0,683	Easy and Feasible
22	0,739	0,578	Easy and Feasible
23	0,783	0,820	Easy and Feasible
24	0,739	0,386	Easy and Medium (no need to revise)
25	0,826	0,462	Very easy and feasible
26	0,783	0,339	Easy and Medium (no need to revise)
27	0,783	0,450	Easy and Feasible
28	0,739	0,386	Easy and Medium (no need to revise)
29	0,652	0,448	Easy and Feasible
30	0,609	0,110	Medium and Bad (not feasible)

Based on the results of item analysis with iteman, it is concluded that out of 30 questions, there are 23 items that are declared feasible and there are 7 items that are declared inappropriate. Then, of the 23 items that were declared feasible, there were 2 questions that needed to be revised. The recapitulation of the results of the analysis is as follows.

Table 6. Recapitulation of ITK dan IDB Feasibility

Number.	Indicator	Total
1	Feasible question	23
2	Question not feasible	7
3	Question difficulty level is very easy	5
4	Question difficulty level is easy	21
5	Question difficulty level is medium	4
6	Question difficulty level is difficult	0
7	Question difficulty level is very difficult	0

Distractor Effectiveness (Item Exemption)

Item analysis also extends to distractor analysis, which is the analysis of test takers' answers to incorrect options (Putri & Rosliyah, 2020). The model assumes that all options must be effective. This means that even if the options are wrong, there must still be a number of test takers who choose them. A good wrong option is one that is able to act as its function, namely as a destroyer, trap, or distractor for some test takers. In order for all options in each item to be effective, the arrangement of the wrong options must be done in such a way that they are not too conspicuous as wrong options. A good wrong option is one that is similar but not the same as the correct option, so that it has a chance of being chosen by careless participants. The results of the distractors or items on this assessment are as follows.

Table 7. Result of Exception Item Functionality Analysis

Question Number	Option	Quantity of Students Who Answered	Prop.	Description
1	A	16	0,696	Answer key
	B	3	0,130	Exception works well
	C	2	0,087	Exception works well
	D	1	0,043	Exception works well
	E	1	0,043	Exception works well
2	A	3	0,130	Exception works well
	B	12	0,522	Answer key
	C	4	0,174	Exception works well
	D	4	0,174	Exception works well
	E	0	0,000	Exception not working
3	A	2	0,087	Exception works well
	B	0	0,000	Exception not working
	C	19	0,826	Answer key
	D	1	0,043	Exception works well
	E	1	0,043	Exception works well
4	A	3	0,130	Exception works well
	B	14	0,609	Answer key
	C	3	0,130	Exception works well
	D	1	0,043	Exception works well
	E	2	0,087	Exception works well
5	A	0	0.000	Exception not working
	B	1	0,043	Exception works well
	C	19	0,826	Answer key
	D	1	0,043	Exception works well
	E	2	0,087	Exception works well
6	A	0	0.000	Exception not working
	B	19	0,826	Answer key
	C	2	0,087	Exception works well
	D	0	0.000	Exception not working
	E	2	0,087	Exception works well
7	A	0	0.000	Exception not working
	B	18	0,783	Answer key
	C	3	0,130	Exception works well
	D	2	0,087	Exception works well
	E	0	0.000	Exception not working
8	A	3	0,130	Exception works well
	B	1	0,043	Exception works well
	C	1	0,043	Exception works well
	D	0	0.000	Exception not working
	E	18	0,783	Answer key
9	A	1	0,043	Exception works well
	B	0	0.000	Exception not working
	C	1	0,043	Exception works well
	D	19	0,826	Answer key
	E	2	0,087	Exception works well

10	A	15	0,652	Answer key
	B	1	0,043	Exception works well
	C	2	0,087	Exception works well
	D	2	0,087	Exception works well
	E	3	0,130	Exception works well
11	A	1	0,043	Exception works well
	B	2	0,087	Exception works well
	C	18	0,783	Answer key
	D	0	0.000	Exception not working
	E	2	0,087	Exception works well
12	A	16	0,696	Answer key
	B	1	0,043	Exception works well
	C	1	0,043	Exception works well
	D	2	0,087	Exception works well
	E	3	0,130	Exception works well
13	A	2	0,087	Exception works well
	B	1	0,043	Exception works well
	C	1	0,043	Exception works well
	D	18	0,783	Answer key
	E	1	0,043	Exception works well
14	A	17	0,739	Answer key
	B	2	0,087	Exception works well
	C	1	0,043	Exception works well
	D	2	0,087	Exception works well
	E	1	0,043	Exception works well
15	A	1	0,043	Exception works well
	B	16	0,696	Answer key
	C	2	0,087	Exception works well
	D	1	0,043	Exception works well
	E	3	0,130	Exception works well
16	A	17	0,739	Answer key
	B	1	0,043	Exception works well
	C	1	0,043	Exception works well
	D	3	0,130	Exception works well
	E	1	0,043	Exception works well
17	A	1	0,043	Exception works well
	B	1	0,043	Exception works well
	C	2	0,087	Exception works well
	D	16	0,696	Answer key
	E	3	0,130	Exception works well
18	A	2	0,087	Exception works well
	B	1	0,043	Exception works well
	C	2	0,087	Exception works well
	D	16	0,696	Answer key
	E	2	0,087	Exception works well
19	A	3	0,130	Exception works well
	B	1	0,043	Exception works well
	C	2	0,087	Exception works well
	D	1	0,043	Exception works well

	E	16	0,696	Answer key
20	A	2	0,087	Exception works well
	B	12	0,522	Answer key
	C	1	0,043	Exception works well
	D	7	0,304	Exception works well
	E	1	0,043	Exception works well
21	A	16	0,696	Answer key
	B	4	0,174	Exception works well
	C	2	0,087	Exception works well
	D	0	0.000	Exception not working
	E	1	0,043	Exception works well
22	A	1	0,043	Exception works well
	B	17	0,739	Answer key
	C	1	0,043	Exception works well
	D	2	0,087	Exception works well
	E	2	0,087	Exception works well
23	A	2	0,087	Exception works well
	B	1	0,043	Exception works well
	C	1	0,043	Exception works well
	D	1	0,043	Exception works well
	E	18	0,783	Answer key
24	A	1	0,043	Exception works well
	B	1	0,043	Exception works well
	C	17	0,739	Answer key
	D	3	0,130	Exception works well
	E	1	0,043	Exception works well
25	A	1	0,043	Exception works well
	B	1	0,043	Exception works well
	C	0	0.000	Exception not working
	D	19	0,826	Answer key
	E	2	0,087	Exception works well
26	A	2	0,087	Exception works well
	B	1	0,043	Exception works well
	C	18	0,783	Answer key
	D	1	0,043	Exception works well
	E	1	0,043	Exception works well
27	A	2	0,087	Exception works well
	B	2	0,087	Exception works well
	C	1	0,043	Exception works well
	D	0	0.000	Exception not working
	E	18	0,783	Answer key
28	A	1	0,043	Exception works well
	B	1	0,043	Exception works well
	C	2	0,087	Exception works well
	D	2	0,087	Exception works well
	E	17	0,739	Answer key
29	A	3	0,130	Exception works well
	B	3	0,130	Exception works well
	C	15	0,652	Answer key

	D	1	0,043	Exception works well
	E	1	0,043	Exception works well
30	A	2	0,087	Exception works well
	B	2	0,087	Exception works well
	C	14	0,609	Answer key
	D	2	0,087	Exception works well
	E	3	0,130	Exception works well

Below are presented examples of problems and the results of calculations on Iteman.

1. Look at the Following Paragraphs for Question 1 and 2!

Culture is the character and identity of every nation. The Indonesian people have an understanding of cultural arts, especially in the form of dances that are spread throughout the country, from Sabang to Merauke, from Miangas to Rote with their unique identity. These dances will be more interesting if they are staged in a kolosan. The objectives of the proposal based on the background experience above are …

A. Socializing regional dance.

B. Preserving regional dance.

C. Demonstrating regional dance.

D. Organize regional dance performances.

E. Increase the love for cultural arts.

From these questions, ITK, ITB, and distractor effectiveness scores were obtained. Below is a graph of the calculation results of the question in question.

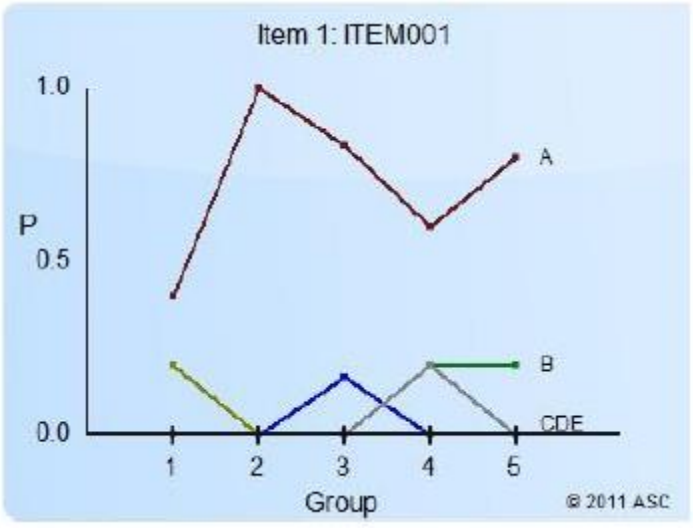

Image 1. Question calculation result

Table of Item Information

Seq.	ID	Key	Scored	Num Options	Domain	Flags
1	ITEM001	A	Yes	5	1	K

Table of Item Statistics

N	P	Total Rpbis	Total Rbis	Alpha w/o
23	0,696	0,139	0,182	0,822

Table of Option Statistics

Option	N	Prop.	Rpbis	Rbis	Mean	SD	Color	
A	16	0,696	0,139	0,182	22,500	4,872	Maroon	**KEY**
B	3	0,130	0,152	0,242	23,000	5,568	Green	
C	2	0,087	-0,251	-0,447	17,000	8,485	Blue	
D	1	0,043	-0,345	-0,763	13,000	0,000	Olive	
E	1	0,043	0,127	0,280	24,000	0,000	Gray	
Omit	0							
Not Admin	0							

Table pf Quantile Plot Data

Option	N	0-20%	20-40%	40-60%	60-80%	80-100%	Color	
A	16	0,400	1,000	0,833	0,600	0,800	Maroon	**KEY**
B	3	0,200	0,000	0,000	0,200	0,200	Green	
C	2	0,200	0,000	0,167	0,000	0,000	Blue	
D	1	0,200	0,000	0,000	0,000	0,000	Olive	
E	1	0,000	0,000	0,000	0,200	0,000	Gray	

Based on the results of the Iteman calculation above, in the statistical item table, it can be seen that the Index of Difficulty Level of the question is 0.696 or question number 1 is in the easy category. Meanwhile, the Distinguishing Power Index is 0.139 or the question is included in the decent category. In addition, in the option statistics table it is known that all distractors were chosen by students. There are 16 learners choosing option A, 3 learners choosing option B, 2 learners choosing option C, 1 learner choosing option D and 1 learner choosing option E. This result is in line with the criteria for determining the effectiveness of distractors, (Dewi, 2015) which states that all distractors must be chosen. Thus all distractor options on the item can be said to be functional.

Conclusion

The Iteman program can be used to conduct item analysis for the End of Semester Examination and other types of school examinations. The results of this study can be an evaluation for the team writing exam questions at the school and utilized by teachers in improving the quality of the items to be tested.

Based on the results obtained in the Iteman program, 30 items tested on students obtained an Alpha Cronbach reliability of 0.820. A test is said to be reliable if the coefficient is 0.60. The Iteman program is expected to continue to develop and be used as a supporter of item analysis, given the importance of the quality of the items tested as an assessment of student learning. Good question items are able to support the expected learning objectives. The role of the teacher as a facilitator and evaluator of learning makes the teacher have to complete item analysis. Research is limited to multiple choice or objective questions, so it is hoped that there will be software development that supports the analysis of assessment questions in the form of descriptions.

References

Amalia, A. N., & Widayati, A. (2012). Analisis Butir Soal Tes Kendali Mutu Kelas Xii Sma Mata Pelajaran Ekonomi Akuntansi Di Kota Yogyakarta Tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*, *10*(1). https://doi.org/10.21831/jpai.v10i1.919.

Arikunto, S. (2018). *Dasar-dasar Evaluasi Pendidikan* (3rd ed.). Bumi Aksara.

Asri, A. S. (2017). Telaah Buku Teks Pegangan Guru Dan Siswa. *RETORIKA: Jurnal Ilmu Bahasa*, *3*(1), 70–32.

Dachliyani, L. (2019). Instrumen yang sahih: sebagai alat ukur keberhasilan suatu evaluasi program diklat (evaluasi pembelajaran). *In MADIKA: Media Informasi Dan Komunikasi Diklat Kepustakawanan*, *1*(1).

Danny, M. (2012). Penjaminan Mutu Satuan Pendidikan Sebagai Upaya Pengendalian Mutu Pendidikan. *Educationist*, *IV*(2), 126–137.

Daryanto. (2013). *Media Pembelajaran Peranannya Sangat Penting Dalam Mencapai Tujuan Pembelajaran*. Gava Media.

Dewi, A. ulfa. (2015). *Analisis Butir Soal Ujian Nasional Mata Pelajaran Ipa Diy an Analysis of the Items of the National Exam of the Natural Science*.

Ilyas, I. (2016). Pendidikan Karakter Melalui Homeschooling. *Journal of Nonformal Education*, *2*(1).

López-Hernández, A., Buckingham, L. R., & Strotmann, B. (2023). Enhancing learning-oriented assessment through co-teaching in higher education. *Studies in Educational Evaluation*, *79*(October), 0–2. https://doi.org/10.1016/j.stueduc.2023.101307.

Nurgiyantoro, B. (2017). Penilaian Pembelajaran Bahasa Berbasis Kompetensi. In *Didaktika: Jurnal Kependidikan* (Vol. 9, Issue 1). https://jurnaldidaktika.org/.

Pradani, R. A., & Efendi, A. (2023). Analisis Butir Soal Ujian Sekolah Menggunakan Program Iteman (Analysis of School Exam Questions Using the Iteman Program). *Indonesian Language Education and Literature*, *8*(2), 275. https://doi.org/10.24235/ileal.v8i2.11002.

Putri, R. A., & Rosliyah, Y. (2020). Analisis Distraktor Butir Soal Bunpou Shokyuu Kouhan Semester 2 Tahun 2018/2019 Prodi Pendidikan Bahasa Jepang UNNES. *Kiryoku*, *4*(1), 18–25. https://doi.org/10.14710/kiryoku.v4i1.18-25.

Rasna, M. F., Ruslau, K. D., & Nur'aini. (2023). Kesiapan Mengajar Mahasiswa Jurusan Pendidikan Matematika Ditinjau Dari Konsep Diri Akademik. *Jurnal Ilmiah Matematika Realistik (JI-MR)*, *4*(1), 63–72. https://jim.teknokrat.ac.id/index.php/pendidikanmatematika/article/view/2627.

Sarker, M. F., & Ullah, M. S. (2023). A review of quality assessment criteria in secondary education with the impact of the COVID-19 pandemic. *Social Sciences and Humanities Open*, *8*(1), 100740. https://doi.org/10.1016/j.ssaho.2023.100740.

Suharjo, Y. U., & Sumardi, V. (2024). Quality Analysis of Indonesian Language Test Questions Class VII SMP Negeri 11 Ruteng Using Iteman Programme Version 4.3. *Formosa Journal of Sustainable Research*, *3*(1), 1–18. https://doi.org/10.55927/fjsr.v3i1.7537.

Copyrights