BioLAE - Journal
Britain International of Linguistics, Arts and Education Sciences Journal
Publish : March - July - November
E-ISSN: 2685 - 4813
P-ISSN: 2685 - 4805
http://biarjournal.com/index.php/biolae

Feasibility Analysis of Items for End of Semester Assesment Questions Class VII Indonesian Subjects of SMPN 1 Ngaglik

Nur Fitriyani[1], Hartono[2]
[1,2]Faculty of Languages, Arts and Culture, Universitas Negeri Yogyakarta, Indonesia
Email: nurfitriyani.2021@student.uny.ac.id,hartono-fbs@uny.ac.id

Abstract: *This research is to determine the feasibility of the test items on the Indonesian language PAS questions for class VII at SMP Negeri 1 Ngaglik for the 2021/2022 academic year usingdocument analysis research with a quantitative descriptive approach. Sources of research data, namely the test questions grid, PAS question sheets in the form of student responses. The research subjects were 32 class VII B students at Ngaglik 1 Public Middle School for the 2021/2022 academic year. The research variable is the feasibility of the Indonesian PAS questions which consist of 40 questions. The research instruments were in the form of documentation guidelines, documentation collection techniques, and data analysis techniques in this study using descriptive statistical techniques. The research results show(1) validity level of 40 questions100% valid, (2) the reliability of the test items was 0.807 and declared reliable, (3) ITK, namely there were 6 easy questions, 30 moderate questions, 3 difficult questions, and 1 very difficult question; IDB there are 27 questions that are appropriate and 13 questions that are not feasible. (4) distractor analysis, there are 30 items (92.5%) that have very good distractor quality and 3 items (7.5%) that have poor distractor quality.*
Keywords: *item analysis; item items; end of semester assessment*

I. Introduction

Assessment is the most important part of the teaching and learning process(Mahardika, 2018). Assessment in the learning process includes activities to collect facts and learning documents of students that can be trusted to make program improvements, if the assessment activities occur as part of a classroom learning program(Supranata & Hatta, 2007). Therefore, assessment functions to help teachers to plan curriculum and teaching in teaching and learning programs, so assessment activities require varied information from each individual and/or group of students and teachers. Teachers also have various strategies and objectives included in the assessment process. It is aligned with(Hoekstra et al., 2023)who argues that teachers have many goals and strategies in their daily work. Educational assessment activities in learning at school are complex activities and involve many aspects and activities in them.(Nurgiyantoro, 2016). Assessment must be carried out and planned properly, so that in order to measure an assessment activity, appropriate measuring instruments are needed and can be accounted for. The teacher's ability to carry out the assessment process is part of the duties and responsibilities in carrying out teaching(Kartini et al., 2022). Sasongko through(Setiawan et al., 2020)argues that assessment is an attempt to obtain various information on a regular, continuous and comprehensive basis about the processes and results of the growth and development that have been achieved by students through learning activity programs.

Learning is a process when someone seeks to find and understand something that is not yet known to become something that is then known (Ashadii & Suhaeb throughEthics Rahmawati & Indriyani Setyaningsih, 2021). Learning outcomes can be used as a determinant of whether the learning process is stopped and continued with the next learning, or the results provide a signal for the learning process to be repeated or improved(Wahyono, 2019).

DOI: *https://doi.org/10.33258/biolae.v5i2.893*

Vygotsky through(Zengilowski et al., 2023)argues that based on the socioconstructivist perspective, learning is a constructive process that is embedded in social and cultural contexts. It is aligned with(Volet et al., 2009)who argued that the emphasis on the learning process as collaborative by students in building knowledge, especially when the classroom encourages the process of community building in a variety of information and discussions. Knowledge and learning are key factors in individual well-being, socio-economic development and international competitiveness(Molnar & Hermann, 2023). Furthermore, optimizing learning behavior and performance is a goal of effective instruction and learning, which has become the interest of educators and researchers alike(Xu et al., 2022).

Assessment of student learning outcomes is carried out using various techniques according to the competencies to be assessed(Subagia & Wiratma, 2016). Assessment of student learning outcomes can be interpreted as an activity of collecting and processing information on student learning outcomes to determine the progress of learning and conclude the results of student learning achievements. In essence, the assessment activities carried out are not solely to assess student learning outcomes, but can also be used to assess the quality of the learning carried out. Assessment of learning outcomes by educational units aims to assess the achievement of graduate competency standards for all subjects by considering the results of student assessments by educators(Setiadi, 2016). One of the assessments of learning outcomes that can be carried out by the teacher is by means of tests or assessments. Assessment is one of the professional competencies that need to be known and through which teachers can obtain educational program issues and reliable data on the status of the education system.(Science & Soromessa, 2015).

The test is one type of measurement tool used to assess student learning outcomes and from the test data will be obtained on the value of learning outcomes(Teluma & Rivaie, 2019).(Ndiung & Jediut, 2020) stated that the test is a number of questions that have right or wrong answers that require answers or responses to measure one's ability. The test is carried out by giving an assignment or series of tasks in the form of questions or orders that must be done by students. A test is an instrument or a systematic procedure for measuring a sample of behavior whose answers are in the form of numbers(Nurgiyantoro, 2016). Suryabrata through(Nani Hanifah, 2014) argued thatgood ice must be reliable, valid, objective, discriminatory, comprehensive, and easy to use. One form of learning outcomes test is a multiple choice test. The multiple choice test is a form of objective test which has the main characteristic of a clear and definite answer key so that the results can be scored objectively(Abdul, 2015).

Compiling test items is not easy, because the questions that will be used for the test should go through an analysis of the feasibility of the items. However, so far there are still many processes that have not been carried out to determine the quality of the instruments to be used for tests, especially PAS. Analysis of the quality of the items is a step that must be taken to determine the degree of quality of a test item, both as a whole and the items that are part of the test item.(Rahayu & Djazari, 2016); (Arifin, 2017). Item analysis is also an examination of test statements, so that specific or detailed information is obtained from the test items prepared.(Himawan & Nurgiyantoro, 2022). In the analysis of the items, it is then possible to calculate the Difficulty Level Index (ITK) and the Difference Power Index (IDB).

Difficulty Level Index (ITK) is an index that shows how easy or difficult an item is for the test takers being tested. The level of difficulty of the items is seen from the ability or ability of students to answer, not from the assumptions of the teacher who composed the questions, because items that are difficult or easy for the teacher are not necessarily difficult or easy for

students(Loka Son, 2019). A good item is one that is of sufficient difficulty level, not too easy or too difficult. Items that are too easy or difficult are just as bad because they cannot reflect the learning outcomes that have been carried out. Analyzing the level of difficulty of the items is to determine the quality of the good questions, knowing the classification of easy, medium, or difficult questions(Abdul, 2015). The ITK of a question item is expressed by an index that ranges from 0.00 to 1.00. An index of 0.00 means that the item in question is very difficult because no one can answer it correctly. Conversely, an index of 1.00 means that the item in question is very easy because all students can answer it correctly. All items are considered feasible if the ITK ranges from 0.15 to 0.85. However, it can be tolerated to be between 0.20 – 0.80.

The validator in this study, namely an Indonesian teacher from North Kalimantan, so that they can share experiences with the validator when discussing online when validating questions for this research. The validator also said that during teaching he had never analyzed item items with the Iteman program, so this research could also help teachers understand the Iteman program to analyze item items. The validator in this study, namely an Indonesian teacher from North Kalimantan, so that they can share experiences with the validator while discussing online when validating questions for this study. The validator also said that during teaching he had never analyzed item items with the Iteman program, so this research could also help teachers understand the Iteman program to analyze item items.

II. Research Method

The type of research used in this research is using document analysis research with a quantitative descriptive approach because the research results are presented in numbers. Quantitative descriptive analysis is used to describe the results of the score acquisition in the item program. The source of the data in this study, namely the question grid related to test questions, PAS question sheets for Class VII Indonesian Language SMP Negeri 1 Ngaglik for the 2021/2022 academic year in the form of student responses. The subjects in this study were 32 class VII B students at Ngaglik 1 Public Middle School for the 2021/2022 academic year. Class VII B was chosen as a sample of the item analysis because this class is a class with heterogeneous academic abilities. Variables in this study, namely the feasibility of the PAS questions with the object of study on the Indonesian language subject test which consists of 40 questions. This is seen from the aspect of the level of validity of the items, the level of difficulty of the items, the functioning of the item distractors, and the level of reliability of the test. The instruments used in this study were documentation guidelines, documentation collection techniques, and data analysis techniques. In this study, descriptive statistical techniques were used, namely percentages.

III. Result and Discussion

The results of the analysis are related to the Odd Semester End Assessment questions for class VII B of SMPN 1 Ngaglik as follows.

3.1 validity

The validity used in this analysis is content validity, namely the validity that must be fulfilled in the test kits, especially the test kits prepared by the teacher to measure the level of student learning success(Nurgiyantoro, 2016). The essence of content validity is determining the accuracy of sampling the teaching materials to be tested. If the test items are clearly intended to measure basic competencies and certain indicators that represent the teaching materials being studied, the test can be said to have good content validity. The procedure

carried out in this assessment, namely making test items based on the grid that has been studied, then the items are reviewed by colleagues. This assessment uses the type of multiple choice questions which total 40 questions. The grid has been reviewed by colleagues, namely Indonesian language teachers based on their suitability with the curriculum used. Meanwhile, in the analysis of the items examined in terms of material, construction, and language.

Table 1. Study Multiple Choice Questions Numbers 1 – 10

spect	**Requirement Type**	**Question Item Number**									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Material	1. The contents of the material are scientifically correct	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	2. There is only one correct answer key	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	3. The content of the material is in accordance with the class/level of education	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	4. Distractions work fine	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
construction	5. The main questions are formulated clearly	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	6. The subject matter does not lead to the correct answer	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	7. Answer choices are clearly formulated	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	8. Homogeneous answer choices	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	9. There is no double negative form	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	10. The length of the answer choices is more or less the same	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	11. The items are not dependent on one another	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	12. Options in the form of numbers/time are sorted	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Language	13. Communicative language	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	14. Grammatical sentences	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	15. Sentences do not have double meaning	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	16. Standard/general/neutral vocabulary	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 2. Study Multiple Choice Questions Numbers 11 – 20

spect	**Requirement Type**	**Question Item Number**									
		11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Material	1. The contents of the material are scientifically correct	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	2. There is only one correct answer key	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	3. The content of the material is in accordance with the class/level of education	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

	4.	Distractions work fine	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
construction	5.	The main questions are formulated clearly	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	6.	The subject matter does not lead to the correct answer	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	7.	Answer choices are clearly formulated	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	8.	Homogeneous answer choices	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	9.	There is no double negative form	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	10.	The length of the answer choices is more or less the same	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	11.	The items are not dependent on one another	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	12.	Options in the form of numbers/time are sorted	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Language	13.	Communicative language	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	14.	Grammatical sentences	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	15.	Sentences do not have double meaning	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	16.	Standard/general/neutral vocabulary	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 3. Study Multiple Choice Questions Numbers 21 – 30

spect	**Requirement Type**		**Question Item Number**									
			21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
Material	1.	The contents of the material are scientifically correct	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	2.	There is only one correct answer key	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	3.	The content of the material is in accordance with the class/level of education	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	4.	Distractions work fine	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
construction	5.	The main questions are formulated clearly	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	6.	The subject matter does not lead to the correct answer	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	7.	Answer choices are clearly formulated	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	8.	Homogeneous answer choices	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	9.	There is no double negative form	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	10.	The length of the answer choices is more or less the same	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	11.	The items are not dependent on one another	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	12.	Options in the form of numbers/time are sorted	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Language	13.	Communicative language	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	14.	Grammatical sentences	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

	15. Sentences do not have double meaning	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	16. Standard/general/neutral vocabulary	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 4. Study Multiple Choice Questions Numbers 31 – 40

Aspect	Requirement Type	Question Item Number									
		31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
Material	1. The contents of the material are scientifically correct	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	2. There is only one correct answer key	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	3. The content of the material is in accordance with the class/level of education	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	4. Distractions work fine	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
construction	5. The main questions are formulated clearly	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	6. The subject matter does not lead to the correct answer	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	7. Answer choices are clearly formulated	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	8. Homogeneous answer choices	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	9. There is no double negative form	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	10. The length of the answer choices is more or less the same	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	11. The items are not dependent on one another	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	12. Options in the form of numbers/time are sorted	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Language	13. Communicative language	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	14. Grammatical sentences	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	15. Sentences do not have double meaning	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	16. Standard/general/neutral vocabulary	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Based on the results of the validity test of 40 PAS questions in class VII Indonesian language at SMP Negeri 1 Ngaglik for the 2021/2022 academic year, it can be concluded that the 40 questions tested are 100% valid questions. It can be seen that the indicators with the questions can be said to be appropriate, so that all the 2021/2022 PAS question items in class VII have a high level of validity.

3.2 Item Analysis

Analysis of the items used in this assessment using Iteman version 4.4. From the 40 questions, data were obtained from the Difficulty Level Index (ITK) and Diversity Index (IDB). ITK is an index that shows how easy or difficult an item is for the test takers being tested. A good item is one that is of sufficient difficulty level, not too easy or too difficult.

Tolerable ITK is in the range of 0.20 – 0.80. The ITK scale classification can be seen in the following table.

Table 5. Difficulty Level Index Classification (ITK)

Difficulty Level Index	
Criteria	**Scale**
0.00 – 0.20	Very difficult
0.21 – 0.40	Difficult
0.41 – 0.60	Currently
0.61 – 0.80	Easy
0.81 – 1.00	Very easy

The differential power of the items is a statement about how much the power of an item can distinguish the ability of participants in the high group and the low group. IDB is an index that shows how much power an item has about ability between participants in the high group and the low group. Theoretically the IDB size can range from -1.00 - +1.00. However, an index that is close to 0 (zero) or negative is declared infeasible. An item is said to be feasible if it has at least an IDB of 0.20. The IDB classification can be seen in the following table.

Table 6. Classification of Difference Power Index

Power Difference Index	
Criteria	**Scale**
0.40 – 1.00	Worthy
0.30 – 0.39	Moderate (Decent and does not need revision)
0.20 – 0.29	Fairly good (Decent and needs to be revised)
Negative – 0.19	Bad (Unworthy)

The results of ITK and IDB calculations in this assessment were not done manually, but were done using a computer program, namely Iteman version 4.4. With the results obtained, it will be known the level of difficulty and feasibility of each item. Each item of PAS Indonesian class VII SMP Negeri 1 Ngaglik for the 2021/2022 academic year, can be seen in the summary table as follows.

Table 7. Results of Item Analysis Based on ITK and IDB

No. Question	**ITK**	**IDB**	**Information**
1	0.438	0.329	Moderate and Decent
2	0.531	0.483	Moderate and Decent
3	0.375	0.105	Difficult and Infeasible
4	0.594	0.282	Moderate and Decent
5	0.500	0.374	Moderate and Decent
6	0.483	0.360	Moderate and Decent
7	0.406	0.406	Moderate and Decent
8	0.531	0.493	Moderate and Decent
9	0.500	0.509	Moderate and Decent

10	0.594	0.059	Moderate and Inadequate
11	0.594	0.251	Moderate and Decent
12	0.500	0.152	Moderate and Inadequate
13	0.531	0.483	Moderate and Decent
14	0.469	0.433	Moderate and Decent
15	0.406	0.216	Moderate and Decent
16	0.563	0.183	Moderate and Inadequate
17	0.656	0.295	Easy and Worthy
18	0.625	0.262	Easy and Worthy
19	0.656	0.295	Easy and Worthy
20	0.531	0.369	Moderate and Decent
21	0.469	0.433	Moderate and Decent
22	0.438	0.360	Moderate and Decent
23	0.500	0.540	Moderate and Decent
24	0.438	0.516	Moderate and Decent
25	0.531	0.369	Moderate and Decent
26	0.750	-0.212	Easy and Infeasible
27	0.531	0.483	Moderate and Decent
28	0.156	-0.496	Very Difficult and Not Worth It
29	0.438	0.318	Moderate and Decent
30	0.531	0.157	Moderate and Inadequate
31	0.781	0.055	Easy and Infeasible
32	0.781	-0.085	Easy and Infeasible
33	0.531	0.227	Moderate and Decent
34	0.531	0.038	Moderate and Inadequate
35	0.531	-0.135	Moderate and Inadequate
36	0.531	0.588	Moderate and Decent
37	0.469	0.149	Moderate and Inadequate
38	0.438	0.116	Moderate and Inadequate
39	0.375	0.611	Hard and Worth it
40	0.406	0.371	Hard and Worth it

Based on the results of the analysis of the items in the Iteman program version 4.4 above, it can be concluded that of the 40 items, there were 27 questions that were feasible and 13 questions that were declared inappropriate. Furthermore, there are 6 questions with easy difficulty level, 30 questions with moderate difficulty level, 3 questions with difficult difficulty level, and 1 question with very difficult difficulty level. The questions that were declared ineligible were due to the ITK or IDB's questions not meeting the standard criteria of ITK or IDB. There is also a recapitulation of the results of the item analysis is as follows.

Table 8. ITK and IDB analysis recapitulation

No.	Indicator	Amount	No. Question Items
1	Decent matter	27	1, 2, 4, 5, 6, 7, 8, 9, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 33, 36, 39, 40
2	Inadequate question	13	3, 10, 12, 16, 26, 28, 30, 31, 32, 34, 35, 37, 38
3	About the level of difficulty is very easy	0	-

4	About the easy difficulty level	6	17, 18, 19, 26, 31, 32
5	About medium difficulty level	30	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 20, 21, 22, 23, 24, 25, 27, 9, 30, 33, 34, 35, 36, 37, 38
6	About the difficulty level is difficult	3	3, 39, 40
7	Problem level of difficulty is very difficult	1	28

3.3 Distractor/Detractor Item Analysis

The distractor serves to identify highly capable test takers. The model assumes that all options must be effective. That is, even if the option is wrong, the number of test takers still has to choose it. There are also distractor results or distracting items in this assessment, as follows.

Table 9. Distractors/Defraud Items

No. Question	Option	Number of Students who Answered	Prop.	Category
1	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	14	0.438	Answer key
	D	8	0.250	The Detractor Works
2	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	17	0.531	Answer key
	D	5	0.156	The Detractor Works
3	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	12	0.375	Answer key
	D	10	0.313	The Detractor Works
4	A	9	0.281	The Detractor Works
	B	19	0.594	Answer key
	C	2	0.063	The Detractor Works
	D	2	0.063	The Detractor Works
5	A	16	0.500	Answer key
	B	5	0.156	The Detractor Works
	C	5	0.156	The Detractor Works
	D	6	0.188	The Detractor Works
6	A	6	0.188	The Detractor Works
	B	14	0.438	Answer key
	C	6	0.188	The Detractor Works
	D	6	0.188	The Detractor Works
7	A	6	0.188	The Detractor Works
	B	6	0.188	The Detractor Works
	C	7	0.219	The Detractor Works
	D	13	0.406	Answer key
8	A	17	0.531	Answer key
	B	5	0.156	The Detractor Works

	C	5	0.156	The Detractor Works
	D	5	0.156	The Detractor Works
9	A	16	0.500	Answer key
	B	5	0.156	The Detractor Works
	C	5	0.156	The Detractor Works
	D	6	0.188	The Detractor Works
10	A	19	0.594	Answer key
	B	3	0.094	The Detractor Works
	C	6	0.188	The Detractor Works
	D	4	0.125	The Detractor Works
11	A	9	0.281	The Detractor Works
	B	19	0.594	Answer key
	C	2	0.063	The Detractor Works
	D	2	0.063	The Detractor Works
12	A	5	0.156	The Detractor Works
	B	6	0.188	The Detractor Works
	C	5	0.156	The Detractor Works
	D	16	0.500	Answer key
13	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	17	0.531	Answer key
	D	5	0.156	The Detractor Works
14	A	5	0.156	The Detractor Works
	B	15	0.469	Answer key
	C	5	0.156	The Detractor Works
	D	7	0.219	The Detractor Works
15	A	7	0.219	The Detractor Works
	B	6	0.188	The Detractor Works
	C	6	0.188	The Detractor Works
	D	13	0.406	Answer key
16	A	5	0.156	The Detractor Works
	B	4	0.125	The Detractor Works
	C	5	0.156	The Detractor Works
	D	18	0.563	Answer key
17	A	21	0.656	Answer key
	B	3	0.094	The Detractor Works
	C	4	0.125	The Detractor Works
	D	4	0.125	The Detractor Works
18	A	5	0.156	The Detractor Works
	B	4	0.125	The Detractor Works
	C	20	0.625	Answer key
	D	3	0.094	The Detractor Works
19	A	21	0.656	Answer key
	B	3	0.094	The Detractor Works
	C	4	0.125	The Detractor Works
	D	4	0.125	The Detractor Works
20	A	17	0.531	Answer key
	B	5	0.156	The Detractor Works
	C	5	0.156	The Detractor Works

	D	5	0.156	The Detractor Works
21	A	5	0.156	The Detractor Works
	B	15	0.469	Answer key
	C	5	0.156	The Detractor Works
	D	7	0.219	The Detractor Works
22	A	6	0.188	The Detractor Works
	B	14	0.438	Answer key
	C	6	0.188	The Detractor Works
	D	6	0.188	The Detractor Works
23	A	5	0.156	The Detractor Works
	B	16	0.500	Answer key
	C	5	0.156	The Detractor Works
	D	6	0.188	The Detractor Works
24	A	6	0.188	The Detractor Works
	B	6	0.188	The Detractor Works
	C	6	0.188	The Detractor Works
	D	14	0.438	Answer key
25	A	17	0.531	Answer key
	B	5	0.156	The Detractor Works
	C	5	0.156	The Detractor Works
	D	5	0.156	The Detractor Works
26	A	24	0.750	Answer key
	B	0	0.000	The Distractor Doesn't Work
	C	0	0.000	The Distractor Doesn't Work
	D	8	0.250	The Detractor Works
27	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	17	0.531	Answer key
	D	5	0.156	The Detractor Works
28	A	17	0.531	The Detractor Works
	B	5	0.156	The Detractor Works
	C	5	0.156	Answer key
	D	5	0.156	The Detractor Works
29	A	14	0.438	Answer key
	B	5	0.156	The Detractor Works
	C	8	0.250	The Detractor Works
	D	5	0.156	The Detractor Works
30	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	17	0.531	Answer key
	D	5	0.156	The Detractor Works
31	A	7	0.219	The Detractor Works
	B	0	0.000	The Distractor Doesn't Work
	C	25	0.781	Answer key
	D	0	0.000	The Distractor Doesn't Work
32	A	0	0.000	The Distractor Doesn't

				Work
	B	25	0.781	Answer key
	C	7	0.219	The Detractor Works
	D	0	0.000	The Distractor Doesn't Work
33	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	5	0.156	The Detractor Works
	D	17	0.531	Answer key
34	A	5	0.156	The Detractor Works
	B	5	0.156	The Detractor Works
	C	5	0.156	The Detractor Works
	D	17	0.531	Answer key
35	A	6	0.188	The Detractor Works
	B	4	0.125	The Detractor Works
	C	17	0.531	Answer key
	D	5	0.156	The Detractor Works
36	A	17	0.531	Answer key
	B	5	0.156	The Detractor Works
	C	5	0.156	The Detractor Works
	D	5	0.156	The Detractor Works
37	A	6	0.188	The Detractor Works
	B	15	0.469	Answer key
	C	6	0.188	The Detractor Works
	D	5	0.156	The Detractor Works
38	A	14	0.438	Answer key
	B	6	0.188	The Detractor Works
	C	6	0.188	The Detractor Works
	D	6	0.188	The Detractor Works
39	A	8	0.250	The Detractor Works
	B	5	0.156	The Detractor Works
	C	12	0.375	Answer key
	D	7	0.219	The Detractor Works
40	A	5	0.156	The Detractor Works
	B	7	0.219	The Detractor Works
	C	13	0.406	Answer key
	D	7	0.219	The Detractor Works

Based on the results of the analysis of the distractor or distractor items, it can be seen that the distractor can be said to function if all the distractors are chosen by the students, while the distractor does not function if no students choose. Decoder item analysis can be said to function properly if at least 5% of the number of test takers is selected. Because the students whose results were analyzed were 32 students, then 5% of 32 is 1.6 and rounded up to 2. Thus, the distractor item is said to function if it is selected by at least 2 students. From the results of an analysis of the 40 PAS questions for class VII Indonesian subject, the following results can be obtained.

Table 10. Results of Distractor/Detractor Item Analysis

Category	No. Question Items	Amount	Percentage
Very good	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 33, 34, 35, 36, 37, 38, 39, 40	37	92.5%
Good	-	0	0%
Not good	26, 31, 32	3	7.5%
Bad	-	0	0%
Very bad	-	0	0%
Amount		40	100%

From the table above, it is known that out of 40 questions there are 30 items (92.5%) that have very good distractor quality because all answer choices are chosen by students and 3 items (7.5%) have less distractor quality. good because in question numbers 26, 31, and 32 there are two answer choices that were not chosen by the test takers at all.

3.4 Reliability

The reliability used in the analysis of this item is Cronbach's Alpha reliability. The technique of calculating the reliability coefficient*Alpha Cronbach is quite widely used for the purposes of testing the results of measurement in the world of education. In this case, the number is calculated as a reliability coefficient to prove the reliability of measuring the results obtained in the study*(Karaca & Sarıgül, 2021). In calculating the index of difficulty level (ITK) and index of difference power (IDB) of item items through the item computer program, the reliability index used is also the reliability coefficient*Alpha Cronbach.*

Itemanprovide statistical data (*statistical scales*) obtained from data entry. From this statistical data will be found information about the magnitude of the reliability coefficient which is calculated by*Alpha Cronbach*, average score, minimum score, maximum score, and average ITK (Mean P) for all items, and IDB average (Mean R) for all items. The analysis for this assessment uses Iteman ver 4.4. The statistical data obtained from Iteman are as follows.

Table 11. Reliability Results

Summary statistics
Statistics, Scored Items
Items:,40,000
Mean:,20.594
SD: , 6671
Min Score:, 10,000
Max Score:,32.000
Mean P:, 0.515
Mean R:, 0.268
Reliability Analysis
Statistics, Scored Items
Alpha:, 0807

Based on statistical data with a total of 40 items tested on 32 students, it can be obtained that the Cronbach Alpha reliability coefficient is 0.807, the average is 20,594, the minimum score is 10, the maximum score is 32, the average ITK for all items is 0.515, and the

average -the average IDB for all items is 0.268. To determine the reliability of the test equipment, it can be seen from the Alpha Cronbach reliability coefficient. Tuckman through Nurgiyantoro (2016) suggests that teacher-made tests are said to be reliable if they have at least a coefficient of 0.60. Based on these assumptions, the test equipment tested in this assessment was declared reliable because the reliability index was greater than 0.60, which was 0.807.

IV. Conclusion

Based on the results of the analysis and discussion above on the End of Semester Assessment questions for class VII Indonesian subjects at SMP Negeri 1 Ngaglik, it can be concluded that (1) the validity level of the item items is 40 itemstested is a question that is 100% valid. It can be seen that the indicators with the questions can be said to be appropriate, so that all the 2021/2022 PAS question items in class VII have a high level of validity. (2) Analysis of the items on ITK, namely there are 6 easy questions, 30 moderate questions, 3 difficult questions, and 1 very difficult item. Furthermore, for the IDB analysis, there are 27 questions that are appropriate and 13 questions that are not feasible. (3) Analysis of distractors or distracting items from 40 items, 30 items (92.5%) had very good distracting qualities and 3 items (7.5%) had poor distracting qualities. (4) The reliability of the test items with a total of 40 items tested on 32 students, the result can be obtained that the Cronbach Alpha reliability coefficient is 0.807, the average is 20,594, the minimum score is 10, the maximum score is 32, the average ITK for all items is 0.515, and the IDB average for all items is 0.268. The test equipment tested in this assessment was declared reliable because the reliability index was greater than 0.60, which was 0.807.

The results of the analysis of the items are in accordance with the research objectives, namely to determine the level of validity of the items, analysis of the level of difficulty and the index of discriminating power of the items, the functioning of the item distractors, and the level of reliability of the test. Thus, the analysis of the items in this study is expected to contribute to conducting further research related to the analysis of item items with the Iteman program. Item analysis needs to be done because this is one of the most important stages in assessing learning by the teacher. If good questions will support teachers and students in achieving learning goals.

References

Abdul, K. (2015). DRAWING AND ANALYZING THE LEARNING OUTCOME TEST Abdul Kadir. Al-Ta'dib, 8(2), 70–81.

Abdullah, MNLY, Omar, AF, Ping, TA, & Chun, TC (2022). Validity and Reliability of the Postgraduate Self-Determined Learning Questionnaire (PSLQ). Asian Journal of University Education, 18(1), 91–105. https://doi.org/10.24191/ajue.v18i1.17175

Arbiatin, E., & Mulabbiyah, M. (2020). Feasibility Analysis of Semester Final Assessment Test Items for Mathematics Class VI at Ampenan 19 Sdn 19 Academic Year 2019/2020. El Midad, 12(2), 146–171. https://doi.org/10.20414/elmidad.v12i2.2627

Arifin, Z. (2017). Learning Evaluation. Bandung: Rosdakarya Youth.

Arvianto, IR (2017). Utilization of the Iteman 3.0 Program for Analysis of Items in the Information and Communication Technology Quiz Competition for Senior High School Equivalents. Respati, 11(33). https://doi.org/10.35842/jtir.v11i33.106

Azwar, S. (1992). Reliability and Validity. Yogyakarta: Student Libraries.

Dwi Saputra, A., Nurul Fauziah, F., Suwandi, S., & Article, S. (2022). Utilization of Indonesian language teaching materials containing local wisdom at SMA Negeri 1 Karanganyar. 8(2), 335–348. http://ejournal.umm.ac.id/index.php/kembara

Ethics Rahmawati, L., & Indriyani Setyaningsih, V. (2021). Students' independent learning in online learning for Indonesian subjects (Students' independent learning in the online learning for Indonesian subjects). Kembara: Scientific Journal of Language, Literature and Teaching, 7(2), 353–365. http://ejournal.umm.ac.id/index.php/kembara

Himawan, R., & Nurgiyantoro, B. (2022). Analysis of exercise items for odd semester final assessment of Indonesian language subjects class VIII SMPN 1 Bambanglipuro Bantul using the ITEMAN program (Analysis of exercise items for odd semester end of semester Indonesian language subjects class. Traveling: Journal of Language Studies, Literature and Teaching, 8(1), 160–180.

Hodiyanto, H., & Saputro, M. (2018). Workshop on making and analyzing items using Iteman at Madrasah Aliyah Miftahul Huda Sungai Ambawang District. Transformation: Journal of Community Service, 14(2), 85–90. https://doi.org/10.20414/transformation.v14i2.578

Hoekstra, NAH, van den Berg, YHM, Lansu, TAM, Mainhard, MT, & Cillessen, AHN (2023). Teachers' goals and strategies for classroom seating arrangements: A qualitative study. Teaching and Teacher Education, 124, 104016. https://doi.org/10.1016/j.tate.2023.104016

Huda, N., & Wahyuni, TS (2020). Use of the Item and Test Analysis (Iteman) Application in the 2019 Middle School IPA Try Out Questions. Science Learning Journal, 4(1), 2527–9157. http://journal2.um.ac.id/index.php/jpsi/article/view/9738

Ibadullah, M., & Maruti, ES (2016). Education Evaluation. Magetan: CV. AE. Graphic Media.

Ida, FF, & Musyarofah, A. (2021). Validity and Reliability in Item Analysis. Al-Mu'Arrib: Journal of Arabic Education, 1(1), 34–44. https://doi.org/10.32923/al-muarrib.v1i1.2100

Iskandar, A., & Rizal, M. (2018). Analysis of the quality of questions in tertiary institutions based on the TAP application. Journal of Educational Research and Evaluation, 22(1), 12–23. https://doi.org/10.21831/pep.v22i1.15609

Karaca, N., & Sarıgül, A. (2021). Aesthetic Perception Scale: Validity and Reliability Study. Shanlax International Journal of Education, 9(S2-Sep), 9–23. https://doi.org/10.34293/education.v9is2-sep.4367

Kartini, A., Eka Sari, D., Youpika, F., Damaianti, V., Supratman Kendang Limun Bengkulu, W., & Article, S. (2022). Development of poetry writing instruments through the android-based "PAP" application. KEMBARA: Scientific Journal of Language, Literature, and Teaching, 8(2), 349–361. http://ejournal.umm.ac.id/index.php/kembara

Loka Son, A. (2019). Instrumentation of Mathematical Problem Solving Ability: Analysis of Reliability, Validity, Difficulty Level and Item Difference Power. Echo of Wiralodra, 10(1), 41–52. https://doi.org/10.31943/gemawiralodra.v10i1.8

Magdalena, I., Fauziah, SN, Faziah, SN, & Nupus, FS (2021). Analysis of Validity, Reliability, Level of Difficulty and Different Power of Items in the Final Semester Examination Theme 7 Class III of Karet 1 Sepatan. STAR: Journal of Education and Science, 3(2), 198–214. https://ejournal.stitpn.ac.id/index.php/ bintang

Mahardika, B. (2018). Application of Portfolio-Based Assessment Methods in Improving Indonesian Language Learning. Elementary: Scientific Journal of Basic Education, 4(1), 33. https://doi.org/10.32332/elementary.v4i1.1030

Molnár, G., & Hermann, Z. (2023). Short- and long-term effects of COVID-related kindergarten and school closures on first-to eighth-grade students' school readiness skills and mathematics, reading and science learning. Learning and Instruction, 83(May 2022). https://doi.org/10.1016/j.learnstruc.2022.101706

Nani Hanifah. (2014). Comparison of Difficulty Levels, Item Distinguishing Power, and Reliability of the Ordinary Multiple Choice and Multiple Choice Tests in the Association of Economics Subjects. SOCIO E-CONS, 6(1), 46.

Ndiung, S., & Jediut, M. (2020). The development of a mathematics learning achievement test instrument for elementary school students is oriented towards higher order thinking. Premiere Educandum: Journal of Basic Education and Learning, 10(1), 94. https://doi.org/10.25273/pe.v10i1.6274

Nurgiyantoro, B. (2016). Competency-Based Language Learning Assessment. Yogyakarta: BPFE.

Nuryanti, S., Masykuri, M., & Susilowati, E. (2018). Iteman analysis and the Rasch model on the development of critical thinking skills instruments for vocational high school students. Journal of Science Education Innovation, 4(2), 224–233. http://journal.uny.ac.id/index.php/jipi

State Officer, HR, & Kurniawati, KRA (2018). CHARACTERISTICS OF TRY OUT OBJECTS IN THE NATIONAL EXAMINATION OF MATHEMATICS HIGH SCHOOLS IN MATARAM CITY USING THE ITEMAN PROGRAM. Journal of VARIAN, 1(2), 11–21. https://doi.org/10.30812/varian.v1i2.67

Putri, RA, & Rosliyah, Y. (2020). Analysis of Item Distractors on Bunpou Shokyuu Kouhan Semester 2 of 2018/2019 Japanese Language Education Study Program UNNES. Kiryoku, 4(1), 18–25. https://doi.org/10.14710/kiryoku.v4i1.18-25

Putri, RH, & Ofianto. (2019). The Effectiveness of Item Analysis Using Anajohn, Anates and Iteman in the Study of USBN Problems in History Lessons in the City of Padang. Student Journal of History and Education, 1(2), 1–11.

Rahayu, R., & Djazari, M. (2016). Analysis of the Quality of Pre-National Exam Questions in Economics and Accounting Subjects. Journal of Indonesian Accounting Education, 14(1). https://doi.org/10.21831/jpai.v14i1.11370

Rotama, AD, Budiutomo, TW, & Bowo, ANA (2020). OBJECT ANALYSIS OF MIDDLE SEMESTER ASSESSMENT PROBLEMS IN CLASS VII PPKN JUNIOR HIGH SCHOOL MUHAMMADIYAH 7 YOGYAKARTA. Academy of Education Journal, 11(01), 24–35. https://doi.org/10.47200/aoej.v11i01.314

Science, N., & Soromessa, T. (2015). The Practice of Student Assessment: The Case of College of. 6(25), 141–150.

Setiadi, H. (2016). IMPLEMENTATION OF ASSESSMENT IN THE 2013 CURRICULUM THE IMPLEMENTATION OF ASSESSMENT IN THE 2013 CURRICULUM Introduction In the 2014/2015 academic year, the 2013 Curriculum has begun to be implemented throughout Indonesia which is a renewal and improvement of Ku. Journal of Educational Research and Evaluation, 20(2), 166–178.

Setiawan, MA, Susongko, P., & Hayati, MN (2020). DIF Detection on the Objective Test Tool for the Science Semester Final Assessment Using the Rasch Modeling. PSEJ (Pancasakti Science Education Journal), 5(2), 23–29.

Volet, S., Summers, M., & Thurman, J. (2009). High-level co-regulation in collaborative learning: How does it emerge and how is it sustained? Learning and Instruction, 19(2), 128–143. https://doi.org/10.1016/j.learninstruc.2008.03.001

Wahyono, H. (2019). Utilization of information technology in the assessment of learning outcomes in the millennial generation in the industrial revolution era 4 . 0. Proceedings of Biology Education, 3(1), 192–201.

Walid, A., & Ramli, M. (2015). Preparation of a High Order Thinking Skills Test Instrument for Class XI High School Students on Reproductive System Material Constructing A Test for Assessing Higher Order Thinking Skills of High School Students on Reproductive System. 371–377.

Xu, Z., Zdravkovic, A., Moreno, M., & Woodruff, E. (2022). Understanding optimal problem-solving in a digital game: The interplay of learner attributes and learning behavior. Computers and Education Open, 3(January), 100117. https://doi.org/10.1016/j.caeo.2022.100117

Zengilowski, A., Lee, J., Gaines, RE, Park, H., Choi, E., & Schallert, DL (2023). The collective classroom "we": The role of students' sense of belonging on their affective, cognitive, and discourse experiences of online and face-to-face discussions. Linguistics and Education, 73, 101142. https://doi.org/10.1016/j.linged.2022.101142