

The Analysis of The Written Test at The End Semester in History Lesson Class X SMA 2 Banda Maluku Tengah Regency

Najirah Amsi

STKIP Hatta-Sjahrir Banda Naira

naj_ira@yahoo.co.id

ABSTRACT

The purpose of this study is to determine: (1) the quality of repetition of the semester subjects History class X SMA Negeri 2 Banda viewed from the aspects of materials, construction and language, (2) the distribution of the level of cognitive sphere of Bloom taxonomy measured in repetition of the end of the semester of the eye lesson of class X history of SMA Negeri 2 Banda, (3) quality of repetition of semester subjects of class X history of SMA Negeri 2 Banda viewed from Validity, Reliability, Distinguishing Power, Level of Difficultness, Key Effectiveness and Effectiveness. This research is descriptive quantitative. Data collection techniques are primary data and secondary data. The results showed that these problems have a relatively high (un-reliable) reliability aspect, which is coefficient - 0.1762. From the logical (logical validity) aspect there is one problem that has not been validated yet. From the problem level aspect, 66.67% of the problems are included in the easy criteria, 30% of the questions including medium criteria and 3.33% of the questions including difficult criteria. From the aspect of the differentiating agent, the grains show 7 or 23.33% have bad distinguishing power, 8 grains or 26,67% have weak distinguishing power, 7 grain or about 23,33% have medium distinguishing power, 6 grain or about 20% classified as having both distinguishing power and 2 grains or 6.67% classified as having excellent distinguishing power.

Keywords: Analysis, Item, End of Semester

INTRODUCTION

The enactment of Curriculum 2013 now provides the freedom for teachers in learning from planning to implementation, including in the preparation of the problem (test) as an evaluation tool. In learning, there are several components that include learning objectives, learning process, and evaluation of learning which a unity that can not be separated is. Evaluation of learning is done to know the achievement of goal so that can know success level of learning which have been done (Ifit Novita Sari, 2015).

Progress of student learning outcomes in a historical learning activity can be illustrated through the evaluation activities. Evaluation of this activity is as a basis for decision making on student success in learning. (D. E. Sari & Wajdi, 2017)

The evaluation of education is (1) the process / activity to determine the progress of education, compared with the intended purpose; (2) efforts to obtain information in the form of feedback (feed back) for the improvement of education. In the process of learning history, assessment of historical learning outcomes is a multilevel work of measurement and assessment relating to: the measurement of historical learning, assessment of historical learning outcomes and inferences of historical learning outcomes.

One of the benchmarks that is often used to describe the success or lack of success in education at all levels of education is the value of the results of the National Exam (UN), because the value of the National Exam (UN) is an indicator that is easily seen by the wider community to be used as a reference about the success of education. (I. N. Sari, 2017)

The execution of national high school exams does not include all subjects. Subjects in the National Exam (UN) majoring in Natural Sciences (IPA) include Indonesian, English, Mathematics, Biology, Physics, and Chemistry. While the majors of Social Sciences (IPS) include Indonesian Language Subject, English, Mathematics, Economics, Geography, and Sociology. So other subjects like History, Civic Education, Religion and others are judged by a school exam. To obtain information on the progress of history subjects is very difficult. Good questions for the semester, semester, and final exam are made by each school or subject teachers' deliberations at the sub-district or district level.

Preparation of questions for SMA Negeri 2 Banda, test questions of semester final tests prepared by the subject history teacher implemented and compiled by the teacher of each subject. The subject matter teacher compiles the question by using a question card to be collected within a certain time. Then the problems are assembled into a package of problems that are distributed to all students of class X SMA Negeri 2 Banda District of Central Maluku.

Implementation of the final examination of the semester in SMA Negeri 2 Banda only preparation of a well-coordinated problem. Furthermore, how the results, whether used already meet the required standards have never held further testing. All is left to the teacher of each subject, so that happens without a deeper idea of how a test should be made, implemented, and analyzed to become a test that qualifies to be tested to a certain standard.

A quality issue is a matter that can provide information as precisely as it suits its purpose. Among the information that is known is to determine which students have or have not mastered the material taught by the teacher.

Matter of class X subjects of class X subjects that tested in SMA Negeri 2 Banda 2016/2017 school year as one of the evaluation tools is made by the subject teachers and previously not tested first, so the quality is not yet qualitatively and quantitatively. The repetition problem as an evaluation tool for learning outcomes that have never been tested need to know the quantity in terms of the attainment of validates requirements, reliability, difficulty and distinguishing power and quality in terms of material, construction, and language. According to Arikunto (2008: 57), the question is said to have good quality when in accordance with the curriculum, fulfil the material, construction and language requirements, have validity, reliability, and high distinguishing ability, moderate level of difficulty and can measure student achievement.

METHODS

The type of this research is descriptive quantitative. It is said to be quantitative because the data in this study is calculated using statistical figures, as well as the end result can be described.

1. Place of Research

The location of the research location is in SMA Negeri 2 Banda Class X Banda District Maluku Tengah Regency Maluku Province.

2. Time of Research

The study time will be conducted in even semesters from April to November 2017 of the 2016/2017 school year.

Population and Sample Research

1. Population Research

Population in this research is all student of class X SMA Negeri 2 Banda Regency of Central Maluku. The total population of 90 people, spread in four parallel classes, namely X1 as many as 25 people, X2 as many as 25 people, X3 as many as 20 people, and X4 as many as 20.

2. Sample Research

According to Arikunto (1997: 107) if the population of more than 100 people, it can be taken between 10-15% or 20-25% or more. However, if the population of fewer than 100 people, then better taken entirely so that the study is a population study. Given the number of X class students in SMA Negeri 2 Banda amounted to less than 100 people, then the sample in the study is 90 people.

Data collection technique

1. Primary data is data obtained from the school or teacher subjects History of SMA Negeri 2 Banda.
2. Secondary data is data obtained from students of class X SMA Negeri 2 Banda Class X1 s.d. X4. This study is ex-post facto, does not manipulate the symptoms studied and the symptoms are naturally present in the field (Kerlinger, 1995: 604). In this study, the student's answer in the answer sheet of History subjects is the answer to the repeat questions of the end of the academic year 2016/2017.

Data analysis technique

Qualitative data

The analysis was done by Gregory Test and multiple choice item selection to obtain the qualitative quality of questions.

To determine the validity of the content is done by means of, namely, the choice of multiple choice items. The results of multiple choice questions will be qualitatively described which includes the conformity of the material with the indicator, constructed with good construction and language conformity with Indonesian rules,

the use of communicative language, not using the prevailing local language taboo, and the choice of answers do not repeat the group of words same.

Quantitative Data

The quantitative analysis is done manually with the help of Excel and Iteman program to test the validity of the item, the reliability of the test, the grain difficulty index, the problem grain problem, and the effectiveness of the distractor every item.

A good cheater can answer more in number to the respondents than the weaker group, and fewer in the clever group (Koyan, 2011: 142).

Test Validity Problem

The technique used to measure the validity of the problem is the product correlation technique of rough numbers. The formula is:

$$r_{xy} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

Information :

r_{xy} = correlation index number r product moment

$\sum xy$ = number of multiplication products between x and y

$\sum x$ = total score of questions (x)

$\sum y$ = total total score (y)

N = total number of samples

Interpretation of the magnitude of the correlation coefficient is as follows:

- 0.80 - 1.00: the validity is very high
- 0.60 - 0.79: High validity
- 0.40 - 0.59: enough validity
- 0.20 - 0.39: low validity
- 0.00 - 0.19: low or invalid validity

(Siregar, 2014: 77)

Test Reliability Problem

Reliable/reliable instruments will produce reliable data as well (Arikunto, 2010: 221).

$$r_{11} = \left[\frac{K}{k - 1} \right] \left[1 - \frac{\Sigma ab^2}{a_1^2} \right]$$

Information :

r_{11} = reliability of the instrument

$K - 1$ = number of items

Σab^2 = number of grain variants

a^2_1 = total variant

Furthermore, in the provision of interpretation of the reliability coefficient of the test (r_{11}) is used the benchmark as follows:

- a. If r_{11} is equal to or greater than 0.70 means the test of learning outcomes being tested for reliability is stated to have a high reliability (= reliable).
- b. If r_{11} is less than 0.70 means the test of learning outcomes being tested for reliability is stated not to have high reliability (un-reliable) (Sudjana, 2011: 209).

Problem Level Problem Test

$$P = \frac{B}{JS}$$

Information :

P = Exchange Index

B = many students who answered correctly

JS = total of all test participant students

Criteria index difficulty

- 0.00 - 0.30 = difficult
- 0.31 - 0.70 = medium
- 0.71 - 1.00 = easy

(Winarni, 2011: 179)

Power Differentiation Test Problem

DP	BA - BB	Atau	DP	2 (BA - BB)
=	$\frac{\quad}{\frac{1}{2} N}$		=	$\frac{\quad}{N}$

Information :

DP = the power of the differentiator

BA = number of correct answers in the top group

BB = number of correct answers in the lower group

N = number of students doing the test

RESULT AND DISCUSSION

Qualitative Data

Problem Analysis of Material, Construction, and Language Aspects

Qualitative analysis is viewed from the material, construction, and language aspects. The method used in this qualitative analysis is panel method, where every material, construction and language expert is given item, study card and study guide. Furthermore, the experts analyzed by themselves. The result of the qualitative analysis is as follows:

Table 3. Problem Analysis Results from Material, Construction and Language Aspects.

No.	Aspects to be reviewed	Number problem which not on criteria	Percentage
	Material		0 %

1	Conformity of questions with indicator	0	
2	Conformity of material inquired with competence of relevance, continuity, high daily wear	0	
3	Problem refers to the cognitive domain	0	
4	There is only one answer key	0	
	Construction		6,67 %
5	The subject matter is formulated briefly, clearly and firmly	0	
6	The formulation of the subject matter and the choice of answers is a necessary statement	0	
7	The subject matter does not give the key clue	0	
8	Principles of free and negative statements	0	
9	The choice of homogeneous and logical answers is reviewed in terms of material	0	
10	Images, graphs, tables, diagrams, or the like are clear and function	0	
11	The length of choice of answers is relatively the same	0	
12	The answer option does not use the "all above wrong / right" statement and the like	0	

13	The choice of answers in the form of numbers/time arranged in order of magnitude of numbers or chronologies	16, 27	10 %
14	The item does not depend on the previous answer	0	
	Language		
15	Use language in accordance with the Indonesian language rules	0	
16	Using a communicative language	0	
17	Not using local language/taboo	0	
18	The choice of answer does not repeat the same word/group of words unless it is a unified whole	3, 19, 30	

Source: primary data processed in 2017

The number of questions not mentioned in the column does not match the criteria means it is in accordance with the criteria specified. Detailed analysis of the material, construction and language aspects are contained in Appendix 3.

a. Distribution Analysis of Cognitive Area Speeches Bloom's Taxonomy

An analysis of the distribution of the Bloom's cognitive domain of Bloom's taxonomy was done by matching the grain with the criteria of Bloom's revised cognitive domain of Bloom's taxonomic cognition.

Table4. Cognitive Data Distribution Database Bloom's Taxonomy

Level	Criteria Bloom's Taxonomy	Number Problem	Percentage
C1 (Recall)	Recognize, identify, recall, retrieve	2, 9, 12, 14, 16, 17, 18, 21, 23,25, 26	36,67 %

	information stored in long-term memory		
C2 (Understanding)	Construct meaning or understanding based on their own initial knowledge, linking new information with existing knowledge, interpreting, modeling, and classifying, summarizing, predicting, drawing conclusions, comparing, matching, and explaining.	1, 3, 4, 5, 7, 10, 13, 15, 19, 20, 22, 24, 28, 29, 30	50%
C3 (Application)	Includes the use of a procedure to solve a problem or perform a task, run, implement.	6,8,11,25,27	16,67 %
C4 (Analysis)	Distinguishing, organizing	-	0 %
C5 (Evaluation)	Checking, criticizing	-	0 %
C6 (Creating)	Create, plan	-	0 %

Quantitative Data

Validity Grain Problem Analysis

Validity test is used to determine the validity of test items. Invalid questions will be discarded and not used while valid questions mean the problem can be used. Criteria if $r_{hitung} > r_{table}$ then the item is valid.

An item can be said to have a high validity or can be declared valid if the scores on the item have correspondence with the total score. In other words, there is a significant positive correlation between the item score with the total score.

Each item that is answered correctly is generally given a score of 1 (one), whereas every wrong answer is given a score of 0 (zero). This type of data is known by the

name of pure discrete data or dichotomic data. While the total score owned by each item is a continuous data.

From the results of the analysis, it turns out that from 30 items tested its validity, only 13 items of them have been declared as a valid problem, that is the number 4, 5, 6, 10, 11, 12, 13, 14, 17, 19, 20, 22, 29. While the other 17 items, namely item number 1, 2, 3, 7, 8, 9, 10, 15, 16, 18, 21, 23, 24, 25, 26, 27, 28, 30 is an invalid item. Data analysis results can be seen in table 5 below:

Table5
Data Result Analysis validity Item Problem Multiple Choice Deuteronomy
Semester Even Subject History Class X SMA Negeri 2 Banda Lesson Year
2016/2017

Criteria	Number	Total	percentage (%)
Valid	4, 5, 6, 10, 11, 12, 13, 14, 17, 19, 20, 22, 29	13	43,33
Invalid	1, 2, 3, 7, 8, 9, 10, 15, 16, 18, 21, 23, 24, 25, 26, 27, 28, 30	17	56,67

Source: secondary data processed in 2017

From the data of the result of the validity of the items of multiple choice test repeat semester subjects History class X Secondary School SMA Negeri 2 Banda 2016/2017 lesson year, it can be seen that some items are valid or validity. Based on the calculation data is known that as many as 13 items about 43.33% or item about the test are declared valid or has validity. While 17 or 56.67% of other items are stated in invalid categories or do not have validity.

Reliability Analysis

A test as a measuring tool can be expressed reliably, if the results of measurements made by using the test repeatedly against the same subject, always show a fixed result, relatively stable or stable.

How to determine the correct test reliability is when done directly to the item items of the test. Calculations performed on the basis of data from instrument test results

alone will result in internal reliability. Internal reliability is obtained by analyzing data from a single test. There are various ways to know the internal reliability, one of which is with the formula K-R. 20.

In search of test reliability the steps that must be done are as follows:

- a Step I: preparing the calculation table in the framework of the test reliability test by displaying 30 items of questions, can be seen in appendix 5.
- b Step II: calculate the total average score, using the formula:

$$X = \frac{\sum X_t}{N} = \frac{1968}{90} = \mathbf{21,87}$$

- c Step III: find the total variant of St2, using the formula:

$$\begin{aligned} \sum \sigma_1^2 &= \frac{\sum X^2}{N} - X^2 \\ &= \frac{43436}{90} - (21,87)^2 \\ \sum \sigma_1^2 &= \frac{43436}{90} - 478,2969 \\ \sum \sigma_1^2 &= 4,32532 \\ \sum \sigma_1^2 &= \mathbf{4,325} \end{aligned}$$

- d Step IV: perform calculations to determine the reliability of the test by using the formula:

$$\begin{aligned} r_{11} &= \frac{K}{K-1} \left(1 - \frac{\sum \sigma_b^2}{\sigma_i^2} \right) \\ r_{11} &= \frac{90}{90-1} \left(1 - \frac{5,079}{4,325} \right) \\ r_{11} &= \frac{90}{89} (-0,1743) \end{aligned}$$

$$r_{11} = \frac{1,011}{2} (-0,1743)$$

$$r_{11} = \underline{\underline{-0,1762}}$$

In the provision of interpretation of the test reliability coefficient (r11) is generally used the following benchmarks:

If the reliability of the test (r11) is equal to or greater than 0.70 means the test of the learning outcomes being tested for its reliability is stated to have a high reliability.

If the test reliability (r11) is less than 0.70 means the test of the learning outcomes being tested for reliability is not considered to have high reliability (un-reliable).

Data reliability analysis results can be seen in the table below.

Table 6

Data Result Analysis Reliability Grain Problem Multiple Choice Deuteronomy Even Semester Subject Lesson History Class X SMA Negeri 2 Banda Lesson Year 2016/2017.

Category	Description
r11 = - 0,1762	Un-reliable

Source: secondary data processed in 2017

Based on the results of the analysis of the reliability of the items of multiple choice test of the even semester of class X history of SMA Negeri 2 Banda in the academic year 2016/2017, has known the magnitude of the reliability coefficient of the test (r11) of -0.1762. Because the reliability of the test (r11) is smaller or less than 0.70 it can be concluded that the double choice test of the even semester test of the class X class history of SMA Negeri 2 Banda 2016/2017 academic year which presents 30 items and followed by 90 students expressed yet has a high reliability (unreliable). To make the instrument reliable then the validity of items that are not good at the waste or replaced the valid instrument must be reliable, but the reliable instrument is not necessarily valid.

Problem Rate Analysis Item Problem

Analyzing grain difficulty level means examining test items or items included in difficult, moderate or easy categories. The items of the test result of the learning result can be expressed as good item items if the item items are not too difficult nor too easy in other words the degree of difficulty of the item is moderate or sufficient. The workable way to find out whether the test result items have a good or unknown degree of difficulty from the size of the grain difficulty index.

Analysis of the results of the results of the repetition of the even semester of the class X history course, for the index of difficulty, is to use the formula:

$$P = \frac{B}{JS}$$

The results of these calculations can be seen in Appendix 6.

From the results of the analysis of 30 items, in the end, it can be seen that as many as 9 (Nine) items are included in the category of items of good quality, in terms of difficulty degree of moderate or moderate matter, that is item number 10, 11, 12, 13, 15, 17, 20, 24 and 25. The items that fall into the category of difficulty amount to 1 (one) item that is number 27. The items that fall into the easy category are 20 (twenty) items, namely point number 1, 2, 3, 4, 5, 6, 7, 8, 9, 14, 16, 18, 19, 21, 22, 23, 26, 28, 29, 30.

From the results of the analysis that has been done, the data and the calculation of the index of difficulty index points, can be seen in table 7 Therefore from it can be obtained information about the level of exchange of items multiple choice test repeat semester even subjects History class X SMA Negeri 2 Banda year lesson 2016/2017 as in table 7 below:

Table 7

**Data Result of Analysis of Tissue Level Item Problem Double Selection
Deuteronomy Even Semester Subject Lesson History Class X SMA Negeri 2
Banda Lesson Year 2016/2017**

No.	Criteria	Number	Total	Percentage
1	Difficult	27	1	3,33 %

2	Average	10, 11,12, 13, 15, 17,20,24,25	9	30 %
3	Easy	1, 2, 3, 4, 5, 6, 7, 8, 9, 14, 16, 18, 19, 21,22, 23, 26, 28, 29, 30	20	66,67 %

Source: secondary data processed in 2017

In relation to the results of problem analysis in terms of difficulty level, then the follow-up that needs to be done by the tester (teacher) is as follows:

- a For items based on the analysis included in the category of good, in the level of difficulty because it is enough or moderate, should the item be immediately recorded in the question bank. Furthermore, the items can be re-issued in the learning tests at the time to come.
- b For items that fall into the category too difficult, there are three possible follow-ups:
 1. The item is removed or dropped and will not be issued in future tests.
 2. Re-examined, traced and traced so that it can be known which factors because the question items are difficult to be answered by the testee (student). After repairs, the items are reissued on the upcoming learning test results.
 3. It should be understood that not every item in the difficult category at any time is taken advantage of which can be used in the classroom grade test. In such circumstances it is appropriate that the items issued by the teacher are items with easy, medium and difficult questions with a ratio of 3: 4: 3 (easy: medium: difficult). Meaning; Category of the easy problem is 3% of all questions, medium category ie 4% of all questions and categories of difficult problems that are 3% of all questions. So that students who have not or little mastered the competency standard and basic competence and get the value below KKM standard (Minimum Criteria of Completion) of each subject need to repeat or not go up the class while students who have mastered the competency standard, basic competence and meet the

minimum mastery criteria (KKM) or have a high ability to level up in the next class.

c For items that fall into the easy category, there is also the possibility of follow-up:

1. The item is removed or dropped and will not be issued in future tests.
2. Researched, tracked and traced carefully to find out the factors that cause the item can be answered by almost all students, there is the possibility of an option or an alternative paired with the items concerned is too easily known by students, where the option is the key answers and which option serves as a distractor or a distraction. Here the teacher should try to fix or replace it with other options, so the key answers with the distraction are difficult to distinguish by students. After the improvement, the question is tried to be issued again on the test of the next learning outcomes.

Distinctive Power Analysis

Analysis of the differentiator is to examine the items in terms of the ability of a problem to distinguish between students who are smart (high-ability) with low-ability students. That is, if the problem is given to children who are clever results show high achievement, and if given to students who are weak results are low.

The test is said to have no distinguishing power if the test, if tested to high achieving students results, is low, but if given to the weaker students the results are higher. And if given to both categories of students, the results are the same. Thus, a test that does not have a differentiating power, will not provide an overview of results that match the actual student's ability.

Knowing the differentiating power of the problem is very important, because one of the foundations that are held to compose the items is the assumption that the ability of the students with one another is different, and the items about the test results should be able to provide test results that reflect there are differences in the abilities that exist among the students.

Analysis of the results of answers from the results of an even semester test of class X history subjects SMA Negeri 2 Banda, test instrument for power differentiator is using the formula:

$$D = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB$$

To determine the number of differentiating power problem, the steps that need to be pursued are as follows:

- To facilitate the calculation of scores contained in the table sorted from the test participants (students) who scored the highest to the lowest score 6.
- Divide or classify the students who numbered 90 people into two groups, namely upper and lower groups. A score of the results of an even semester test subjects of class X history of SMA Negeri 2 Banda 2016/2017 academic year.
- Provide codes for student groupings of the two categories
- Finding or calculating BA, JA, PA, BB, JB and PB. Calculations for obtaining BA, JA, PA, BB, JB and PB.
- Finding or calculating the number of differentiating power problem for 30 items of rehearsal of even semester subjects of class X history of SMA Negeri 2 Banda academic year 2016/2017.

Table 8.

Category of Different Levels in Different Power

No	Question Proportion of upper group participants who answered correctly	Proportion of lower group participants who answered correctly	Distinguishing Power	Description
1	0,95	0,67	0,28	Problem fixed
2	0,92	0,67	0,25	Problem fixed
3	0,99	1	-0,01	Problems rejected

4	0,88	0,33	0,55	Good question
5	0,84	0,67	0,17	Problems rejected
6	0,74	0,00	0,74	Good question
7	0,77	1	-0,23	Problems rejected
8	0,76	1	-0,24	Problems rejected
9	0,83	1	-0,17	Problems rejected
10	0,68	0,00	0,68	Good question
11	0,69	0,33	0,36	Accept and fix
12	0,63	0,00	0,63	Good question
13	0,67	0,00	0,67	Good question
14	0,82	0,00	0,82	Good question
15	0,66	0,33	0,33	Accept and fix
16	0,71	0,33	0,38	Accept and fix
17	0,52	0,33	0,19	Problems rejected
18	0,85	1	-0,15	Problems rejected
19	0,74	1	-0,26	Problems rejected
20	0,51	0,33	0,18	Problems rejected
21	0,86	0,67	0,19	Problems rejected
22	0,83	0,33	0,50	Good question
23	0,87	1	-0,13	Problems rejected
24	0,40	0,00	0,40	Good question
25	0,47	0,33	0,14	Problems rejected
26	0,85	0,33	0,52	Good question
27	0,08	0,00	0,08	Problems rejected

28	0,98	0,67	0,31	Accept and fix
29	0,80	0,67	0,13	Problems rejected
30	0,80	0,67	0,13	Problems rejected

Source: Secondary data processed in 2017

Based on the data in the above table it can be seen that 15 points or 50% of the rejected questions are 3.5, 7, 8, 9, 17, 18, 19, 20, 21, 23, 25, 27, 29, and 30 categories of questions rejected, 2 items or 6.67% category of problems corrected, 4 items or 13.33% category of questions received and improved, and 9 points or 30% category of good questions.

Sourced from the data presented can be obtained the information as listed in table 9 below:

Table 9

**Data Result Analysis Differentiation Power Item Reality Problem Event
Semester Subject Subject History Class X SMA Negeri 2 Banda Lesson Year
2016/2017**

No.	Differentiating power question	Number of questions	Total of test	Percentage
1	Poor	3, 7, 8, 9, 18, 19, and 23	7	23,33 %
2	Weak	5, 17, 20, 21, 25, 27, 29 and 30	8	26,67 %
3	Medium	1, 2, 11, 15, 16, 24 and 28	7	23,33 %
4	Good	4, 10, 12, 13, 22, and 26	6	20 %
5	Very good	6 and 14	2	6,67 %

Source: Secondary data processed in 2017

It can be seen that from as many as 30 items that are issued in the repetition of the semester of class X history of SMA Negeri 2 Banda of the academic year 2016/2017, the points of the matter when viewed from the differentiating grains indicate 7 grains or 23.33% (8) or about 23.33% have medium distinguishing power, 6 grains

or about 20% are classified as having good distinguishing properties and 2 grains or 6.67% are classified as having excellent distinguishing power. So from the point that the problem is only about 15 grains or 50% has adequate distinguishing power.

Follow-up on the results of analyzing the differentiating power of the test results are:

1. Question items that have a differentiating power of a good question to be entered or recorded in the bank book about. The items on the upcoming learning test result can be excluded because the quality is sufficient.
2. Item of the problem that its distinguishing power is still low (poor), there is two possible follow-ups are:
 - a Retrieved later repaired, and after repair can be submitted in the upcoming learning test results. Later the problem is analyzed again whether its differentiating power increase or not.
 - b Removed or dropped and for future tests, the item will not be issued again
3. Special items that the number of discrimination index because of the negative sign, preferably on the test results of learning to come no longer need to be issued because the item is so poor quality.

Function Analysis Key Answer and Distractors

Analyzing the distractor function is often referred to as analyzing the dispersion pattern of items. As for the pattern of the distribution of items, the students may decide the answers of the answers to the possibility of options that have been paired on each item. This is intended to find out whether or not the answer is available.

The distractor is said to be at least 5% of all test participants (students).

Based on the analysis of the calculation of the number of test participants (students) who chose the option / alternative answers to the test on the repetition of the event semester of class X history of SMA Negeri 2 Banda in the 2016/2017 academic year, the distribution of data can be seen in table 10 below:

Table 10
Key Effectiveness and Distractor Data (spotters)

No of Test	Problem Answer					Key	Description	
	A	B	C	D	E		Detector/Distractor	Key
1	E	E	BE	BE	BE	B	Not functioning	Functioning
2	BE	E	E	BE	BE	C	Not functioning	Functioning
3	E	BE	BE	BE	BE	A	Not functioning	Functioning
4	E	BE	E	BE	BE	C	Not functioning	Functioning
5	E	BE	E	BE	BE	A	Not functioning	Functioning
6	E	E	E	E	E	B	Functioning	Functioning
7	E	E	BE	BE	BE	A	Not functioning	Functioning
8	E	BE	E	BE	E	A	Not functioning	Functioning
9	E	E	BE	BE	BE	A	Not functioning	Functioning
10	E	E	E	E	E	C	Functioning	Functioning
11	E	E	E	E	E	E	Functioning	Functioning
12	E	E	E	E	E	C	Functioning	Functioning
13	E	E	E	E	E	C	Functioning	Functioning
14	BE	E	E	E	E	D	Not functioning	Functioning
15	E	E	E	E	E	B	Functioning	Functioning
16	E	E	E	E	E	D	Functioning	Functioning
17	E	E	E	E	E	E	Functioning	Functioning
18	E	E	E	BE	BE	B	Not functioning	Functioning
19	E	E	E	E	E	C	Functioning	Functioning
20	E	E	E	E	E	C	Functioning	Functioning
21	E	E	BE	BE	BE	B	Not functioning	Functioning
22	E	E	BE	BE	BE	A	Not functioning	Functioning
23	E	E	BE	BE	BE	A	Not functioning	Functioning
24	E	E	E	E	E	C	Functioning	Functioning
25	E	E	E	E	E	C	Functioning	Functioning
26	BE	BE	E	E	BE	D	Not functioning	Functioning
27	BE	E	E	E	E	A	Not functioning	Not functioning
28	BE	BE	E	BE	BE	C	Not functioning	Functioning
29	BE	E	BE	BE	E	E	Not functioning	Functioning
30	E	BE	E	BE	BE	A	Not functioning	Functioning

Source: Secondary data is processed in 2017

Information:

- E: Effective
- BE: Not Effective

Then comes from the data presented it can be obtained the information as listed in table 11 below.

Table 11
Data about Functioning No Key Answers in an Even Semester Reexamination
Subject Subject History Class X SMA Negeri 2 Banda Lesson 2016/2017

No.	Effectiveness Key Answer	Total	Percentage
1	Has functioned properly	29	96,67 %
2	Not functioning properly	1	3,33 %

Source: secondary data processed in 2017

Based on the calculation analysis in the attachment and table about the effectiveness of the answer keys can be known there are 29 items or 96.67% has functioned well and that does not work only 1 item or 3.33%.

To know each item on the problem of multiple choices, function or not distractor based on the analysis of the calculation of the number of students who choose the option / alternative answers to the test questions of the whole item can be seen in table 12 below.

Table 12
Data about Functioning No Distractor (Pengecoh) in OUT Even Anthonic Test
Subject Subject History Class X SMA Negeri 2 Banda Lesson 2016/2017

No.	Conditions Distractors/outside	Total	Percentage
1	Telah berfungsi dengan baik	12	40 %
2	Tidak berfungsi dengan baik	18	60 %

Source: secondary data processed in 2017

Based on the calculation analysis in the appendix and the table about the condition of distractor can be seen that the average test of multiple choice of subjects of history has not functioned properly. This is because about 60% of the distractor options chosen by the test takers (students) are not functioning properly and only 40% of the distractor options selected by the student test takers have functioned properly. This means that the existing distractor has not been able to stimulate or deceive the

test participants (students) who follow the test to choose which is not the key answer or distractor. As a follow up to the results of analyzing the distractors that have not functioned properly should be repaired or replaced with another distractor.

CONCLUSION

Based on the results of the analysis of repetition of the semester subject's class X class history of SMA Negeri Banda 2016/2017 lesson year can be concluded that:

1. In terms of material, construction and language, 0% do not meet the material aspect, 6.67% does not meet the construction aspect and 10% does not meet the language aspect.
2. Distribution of cognitive level of Bloom's measured cognitive domain is C1 43,33%, C2 56,67%, C3 3,33%, C4 0%, C5 0% and C6 0%.

From the logical (logical validity) aspect there is one invalid problem. From reliability aspect obtained reliability coefficient - 0,1762 with criteria not yet have high reliability (unreliable). From the problem level aspect, 66.67% of the problems are included in the easy criteria, 30% of the questions including medium criteria and 3.33% of the questions including difficult criteria. From the aspect of the differentiating power of the item 15 grains or 50% of the rejected matter, 2 items or 6.67% of the problem category were corrected, 4 items or 13.33% of the accepted and improved categories, and 9 points or 30% of the category of good. Grains of the problem when viewed from the differentiation of the grains indicate as much as 7 grains or 23.33% have an ugly distinguishing power, 8 grains or 26.67% have a weak distinguishing power, 7 grains or about 23.33% have medium distinguishing power, 6 grains or about 20% are classified as having good distinguishing properties and 2 grains or 6.67% are considered to have excellent differentiating power. So from the point that the problem is only about 15 grains or 50% have adequate distinguishing power. From the effectiveness of the key answers can be found there are 29 items or 96.67% has functioned well and that does not work only 1 item or 3, 33%. From the effectiveness aspect, there are 18 problems or 60% option distractor chosen by

the test participant (student) not yet functioning properly and only 12 problems or 40% option distractor selected by student test (student) has functioned well.

REFERENCES

- Anas Sudijono, 1995. *Pengantar Evaluasi Pendidikan*. Jakarta : Rajawali Pers
- Anastasi, Anne dan Urbina, Susana. 1997. *Psicohological Testing*. (Seventh Edition). New Jersey: Prentice-Hall, Inc.
- Arikunto, S.. 2005. *Manajemen Penelitian*. Jakarta : Rineka Cipta.
- Candiasa, I Made. 2004. *Analisis Butir*. Singaraja : IKIP Negeri Singaraja
- Candiasa, I Made. 2010. *Pengujian Instrumen Penelitian Disertai aplikasi ITEMAN dan BIGSTEPS*. Singaraja: Unit Penerbitan Universitas Pendidikan Ganesha.
- Depdiknas. 2008. *Panduan Analisis Butir Soal*. Jakarta : Dirjen Manajemen Pendidikan Dasar dan Menengah.
- Depdiknas. 2008, *Panduan Penulisan Butir Soal*. Jakarta: Dirjen Manajemen Pendidikan Dasar dan Menengah.
- Depdiknas. 2007. *Model Pembelajaran Sejarah SMA*. Jakarta: Balitbang Pusat Kurikulum.
- Depdiknas. 2007. *Naskah Akademik Kajian Kebijakan Kurikulum Mata Pelajaran Sejarah*. Jakarta: Balitbang Pusat Kurikulum
- Kerlinger, Fred N. 1995. *Asas-asas Penelitian Behavioral*. Yogyakarta: Gadjah mada University Press.
- Koyan, I Wayan, 2011. *Asesmen Dalam Pendidikan*. Singaraja: Universitas Pendidikan Ganesha
- Safari. 2003. *Evaluasi Pembelajaran*. Jakarta : Depdiknas
- Saifudin, Azwar. 1996. *Reliabilitas dan Validitas Edisi Ketiga*. Yogyakarta: Pustaka Pelajar
- Saifudin, Azwar. 2006. *Tes Prestasi Fungsi dan Pengembangan Pengukuran Prestasi Belajar*. Yogyakarta: Pustaka Pelajar.
- Surapranata, S. 2004. *Analisis Validitas, Reliabilitas dan Interpretasi Hasil Tes: Implementasi Kurikulum 2004*. Bandung : Remaja Rosdakary

- Suharsimi, Arikunto. 2008. *Dasar-dasar Evaluasi Pendidikan Edisi Revisi*. Jakarta: Bumi Aksara
- Suryasubrata, Sumadi. 2005. *Alat Ukur Psikologis*. Yogyakarta : Andi
- Suryabrata, Sumadi. 1984. *Psikologi Pendidikan*. Jakarta : Raja Grafindo Persada.
- Undiksha. 2006. *Kumpulan Abstrak Penelitian*. Singaraja : Program Pasca Sarjana Undiksha.