

(UNAS). So, analyzing MCQ test is importance to do by the teacher to know the quality of their test.

Considering the explanation above, the writer chooses *An analysis of English Multiple Choice Questions (MCQ) Test of 7th grade at SMP BUANA Waru Sidoarjo*. The researcher wants to measure the quality of English Multiple Choice Question (MCQ) test that has been constructed by English teacher in SMP BUANA Waru Sidoarjo.

By doing this research, the writer hopes the teacher can construct a good MCQ test. So the result of the test can be accountable to measure the students' success during learning process in the class. Moreover, by giving MCQ test to their students, the teacher starts to prepare them to do Final Examination (UNAS), because Indonesia still uses MCQ model in UNAS.

This research ever does by some researchers before, the first is Aan Kulistyono Ujiwandi, his research was about *An Analysis Of The Teacher Made English Test Items In UAS Semester 2 2005 / 2006 for the first year students of SMAN 1 Puri Mojokerto*. His research focused on the content validity, reliability, analysis of item difficulty and index of discrimination. The differences of the writer study with this research are the writer just focused on MCQ test and the writer add face validity to be analyzed and measure the quality of the MCQ English test based on the criteria of constructing a good English MCQ test.

The second is Sigit Eko Saputro, his research was about *An Analysis of The Teacher Made Of The First Year Student Of SMA 2 Magetan*. His research almost same with the first research, he focused on the content validity, reliability, index of difficulty, and index of discrimination. The differences of the writer study with this research are the writer more specific in MCQ test, and the writer try to measure the validity include face validity and content validity which did not analyze by the researcher before and the writer also add an analysis of distractor in her research.

The third is Millatul Islamiyah, her research was about *Content Validity and Item Analysis Of Semester II English Final Test For Tenth Grade Students Of SMAN 3 Sidoarjo*. Her research was focused on content validity and item analysis only. The differences this research with the writer's research is lied on the writer focused on the English MCQ test that has been written by English teachers in SMP Buana Waru, so the result of this research can be used as an instrument to measure their quality in constructing English MCQ test, but the previous research did not mention who was the test maker. So her result did not give significant English system learning in this school especially for the teacher.

The fourth is Ria Dewi Pratiktasari, her research with the title *An Analysis of Semester II English Summative Test for the Eighth Year Students of SMPN 1 Slahung Ponorogo*. Her research focused on content validity and item analysis includes index of difficulty and index of discrimination. The difference of this research with the writer's research is the writer more focuses on MCQ test. The

an obvious appeal, since the test only contains what it is thought that the students have actually encountered, and thus can be considered, in this respect at least, a fair test. The disadvantage is that if the syllabus is badly designed, or the books and other materials are badly chosen, the results of a test can be very misleading. Successful performance on the test may not truly indicate successful achievement of course objectives.

The alternative approach is to base the test content directly on the objectives of the course. This has a number of advantages. First, it compels course designers to be explicit about objectives. Secondly, it makes it possible for performance on the test to show just how far students have achieved those objectives.

Progress achievement tests, as their name suggests, are intended to measure the progress that students are making. They contribute to formative assessment. Since 'progress' is towards the achievement of course objectives, these tests, too, should relate to objectives. But how? One way of measuring progress would be repeatedly to administer final achievement tests, the (hopefully) increasing scores indicating the progress made. This is not really feasible, particularly in the early stages of a course. The low scores obtained would be discouraging to students and quite possibly to their teachers. The alternative is to establish a series of well-defined short-term objectives. These should make a clear progression towards the final achievement test based on

the text, it can help students to efficient the time. The teacher also consistent in using the capital letter, but there are some question which has not efficient stems, for example in question number 6, 12, 43, and 48. There are some repetition words and it can waste the students' time to do the test.

b) Analyzing of Content Validity

Table analysis in table 2 appendix II show us the syllabus of English subject in 7th grade SMP Buana Waru included standard of competencies and basic competencies of 2010 curriculum is used. Then, to categorize whether the English MCQ test of SMP Buana Waru fulfill the agreement of content validity or not, the *standar isi* table is matched with the materials in the test. Furthermore, place each item number in appropriate intersection of *standar isi* table of 2010 curriculum to identify the representative sample.

Table analysis in appendix 3 is used to ease the writer in analyzing content validity of English Multiple Choice (MCQ) test that have been constructed by English teacher in 7th grade in SMP BUANA Waru. There are six columns in that table. The first column contains standard competence, second column contains of basic competencies, the third column contains of indicators, the fourth column contains of learning experience, the fifth column contains of items test that appropriate with the basic competencies, the next column contains of the number of items test (S) and the last column contains

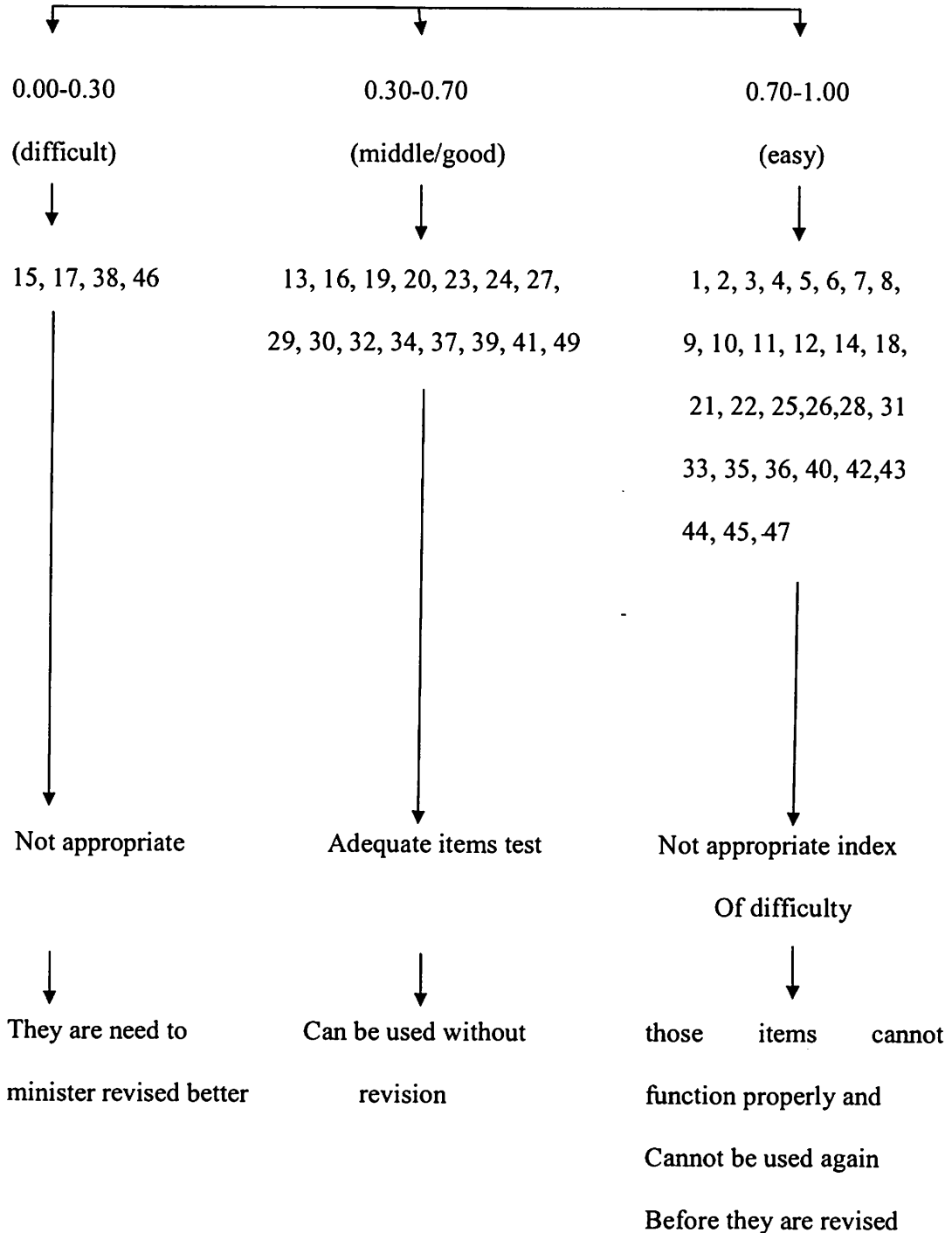
the percentage of total numbers of particular items represent the related basic competence.

According to J. B Heaton, the test can be said has a good content validity if it covers all the contents as stated in the curriculum. Based on the result of analyzing content validity in appendix 3, the percentage of every aspect of the learning content is conclude as follows:

1. There are 8% or 4 items for speaking which focus on asking and giving services.
2. There are 4% or 2 items for speaking which focus on like and dislike.
3. There are 4% or 2 items for speaking which focus on giving response (sorry, thanks, please)
4. There are 6% or 3 items for speaking which focus on greetings and giving response.
5. There are 20% or 10 items for listening which focus on giving opinions and answer the question based on the text.
6. There are 12% or 6 items for reading which focus on announcement.
7. There are 30% or 15 items for reading which focus on descriptive and procedure text, with 16% or 8 items for procedure text and 14% or 7 items for descriptive text.
8. There are also 16% or 8 items test that did not cover the available materials.

2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 18, 21, 22, 25, 26, 28, 31, 33, 35, 36, 40, 42, 43, 44, 45 and 47, so the totals are 31 items. Moreover there are 8% of the tests that categorized as difficult item tests because just there are few students who can answer these items correctly with index of difficulty value 0,00 – 0.30. The item numbers of the test which can be categorized as difficult item test are the item test of numbers 15, 17, 38 and 46, so the totals are 4 items. Those items are needed to ministere revised better because they are not able to do well by the students. And the other 30% is categorized as middle or good item tests because has index of difficulty value between 0.30 – 0.70, they are the item tests of numbers 13, 16, 19, 20, 23, 24, 27, 29, 30, 32, 34, 37, 39, 41 and 49, so the totals are 15 items. Those items are concluded as adequate item test. it means those items can be fully utilized and become reference to design other items for other test. We can see the conclusion of index of difficulty in 7B by this following diagram.

Diagram I: Diagram of analyzing index of difficulty for 7B



middle or good item tests because has index of difficulty value between 0.30 – 0.70, they are the item tests of numbers 1, 2, 5, 7, 8, 9, 11, 13, 15, 16, 17, 18, 19, 21, 23, 24, 25, 27, 29, 30, 32, 34, 35, 38, 39, 40, 46 and 47, so the totals are 28 items. Those items are concluded as adequate item test. It means those items can rewrite without revision. We can see the conclusion of index of difficulty in 7B by this following diagram.

Diagram II: Diagram of analyzing index of difficulty for 7D

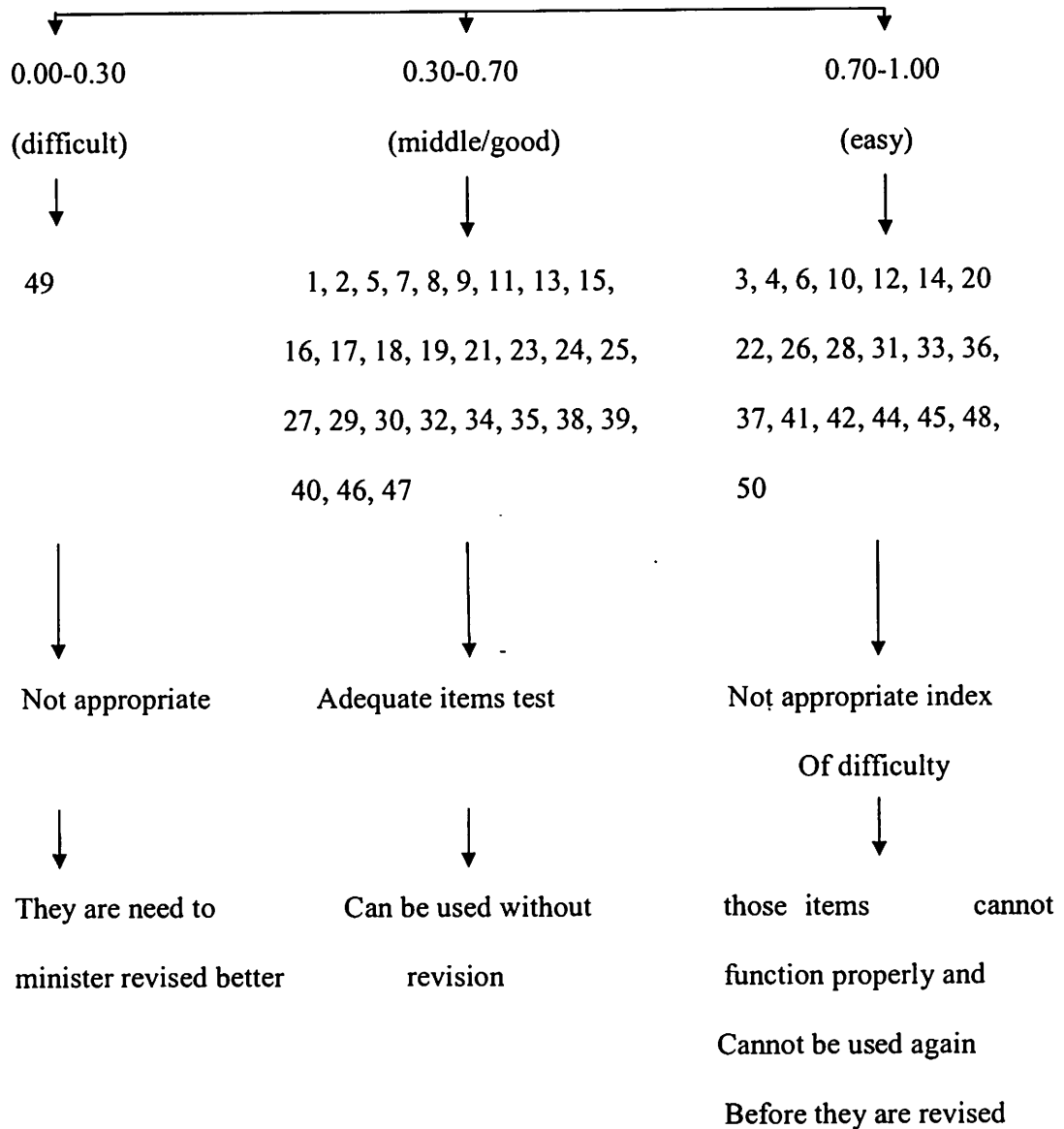
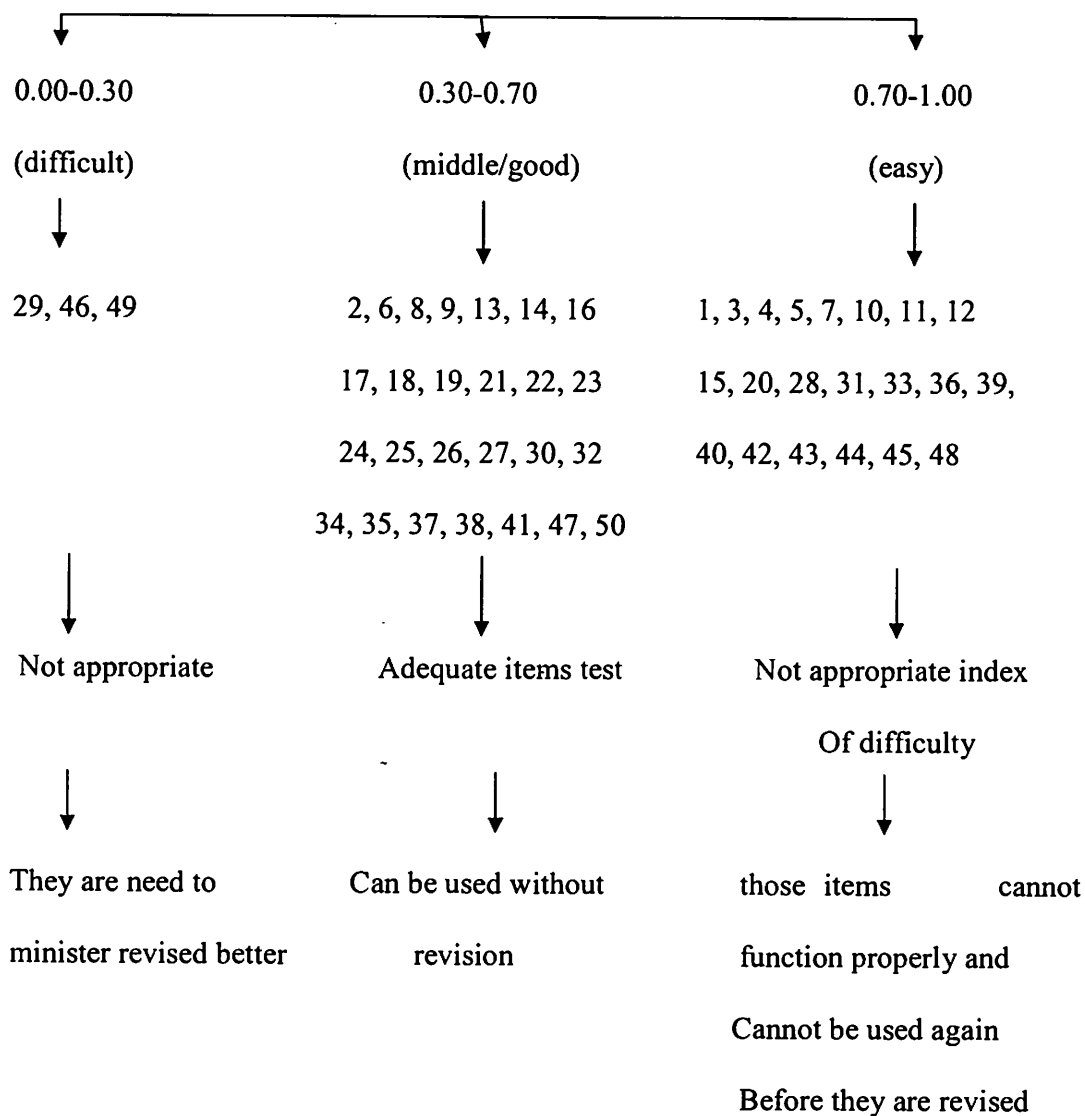


Diagram III: Diagram of analyzing index of difficulty for 7E



Based on those result we can generally conclude that index of difficulty in SMP BUANA Waru are good with index of difficulty value 0.30-0.70. Although the result in 7B shown that most of items categorized as easy item test with 62%. Those

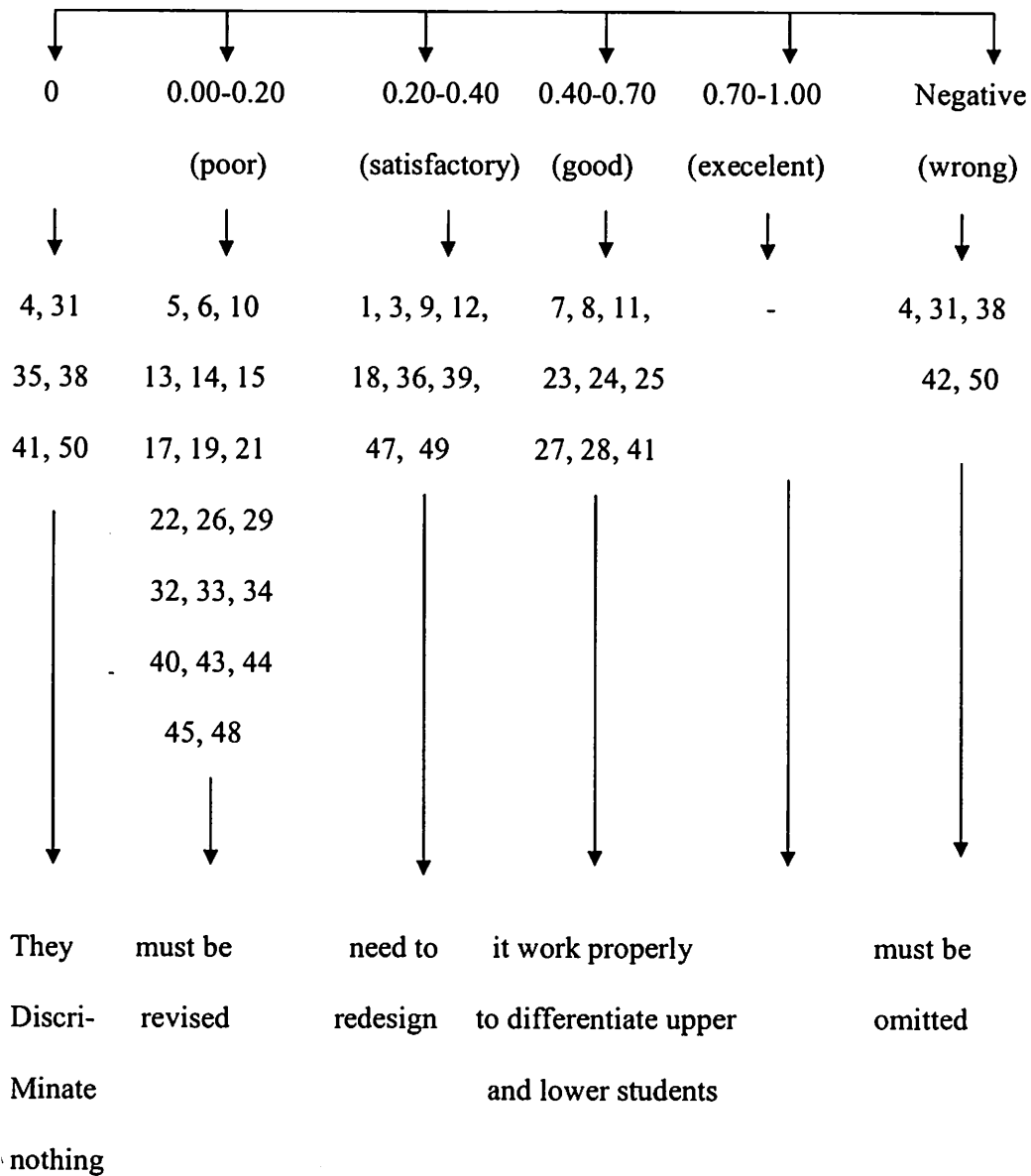
which have discrimination value between 0,20 – 0,40 are number 1, 3, 9, 12, 18, 36, 39, 47, and 49.

From the appendix 6, we can also see that there are 10 items number or around 20% all of the test have the discrimination value between 0,40 – 0,70 and it is categorized as a good item test. It means these items are work properly to differentiate upper students from lower students and they could be utilized for the other test. the items number which have discrimination value between 0,40 – 0,70 are number 7, 8, 11, 23, 24, 25, 27, 28, and 41.

In other hand, there are 4 items or around 8% which have negative discrimination value and it is categorized as wrong item test, they are number 4, 31, 35, 38, 42, and 50. Based on Arikunto, the items of the test which have negative discrimination value must be omitted. It means these items discriminate entirely in wrong way and need to be revised, it can make the purpose to discrimination test do not reach, it is to determine between upper students and lower students.

There are 6 items or 12% which have 0,00 discrimination value, it means these items did not work properly because they discriminate nothing. They are items number 4, 31, 35, 38, 42 and 50. To make it clearly, see this diagram following:

Diagram 4: diagram of analyzing index of discrimination of 7B



b) Analyzing of Index of Discrimination of 7D

In appendix 7 we can see the result of calculating of index of discrimination of 7D. We can conclude that the big number of items tests in 7D are 24 items or 48% which have discrimination value between 0,20 – 0,40, it is categorized as satisfactory item test. It means they still need to be redesigned because they are doubtful to use in the other test. The items test which have discrimination value between 0,20 – 0,40 are number 1, 3, 5, 6,8, 10, 12, 13, 15, 16, 21, 22, 25, 28, 34, 36, 37, 42, 43, 44, 45, 47, 48 and 50.

The second, there are 18 items number or around 36% all of the test have the discrimination value between 0,40 – 0,70 and it is categorized as a good item test. It means these items are work properly to differentiate upper students from lower students and they could be utilized for the other test. The items number which have discrimination value between 0,40 – 0,70 are number 2,7, 9,11,17,18;19, 20, 23, 24, 27, 29, 30, 31, 32, 35, 41 and 46.

In other hand, there are 2 items or around 4% which have negative discrimination value and it is categorized as wrong item test, they are number 33 and 39. Based on Arikunto, the items of the test which have negative discrimination value must be omitted. It means these items discriminate entirely in wrong way and need to be revised, it can make the

can understand the test easily. Second, this test only one feature at a time should be tested. Third, each option also has grammatical correct when placed in stem, so the stems not give clues or question to the question. Furthermore, it can be categorized that the distractors is useful because it seems correctly. Fourth between a question with another questions be dependent and does not give clues to other question. And the last all multiple choice items should be at a level appropriate to the proficiency level of the testees.

From the result of analysis the validity include face validity or the performance of the test and content validity, it is conducted that English Multiple Choice Questions (MCQ) test that has been constructed by an English teacher of 7th grade in SMP BUANA Waru Sidoarjo has good face validity. This test has a good layout with the good arranging. The teacher divided the paper test into 2 columns, it has some advantages. Firstly, it is no spent large place. And the secondly, the students more easy to read the text, it can help students to efficient the time. The teacher also consistent in using the capital letter, but there are some question which has not efficient stems, for example in question number 6, 12, 43, and 48. There are some repetition words and it can waste the students' time to do the test. Furthermore, from the result of analyzing content validity, we can conclude English Multiple Choice Question (MCQ) test that has been constructed by English teacher in 7th grade in SMP BUANA Waru is good since 84% items test represent all materials. It is more than 50%, according to Bloom if the agreement of the test is 50% or more, it can be

conclude that the test has high content validity. Moreover, there are 8 items or 16% all of the test did not cover the materials, they are the items test number 18, 25, 34, 35, 36, 39, 46 and 47. Those items are not suitable with the indicators of standard and basic competencies and were not taught in this semester.

From the result of analyzing of reliability, we can conclude that English Multiple Choice Questions (MCQ) test that has been constructed by an English teacher of 7th grade in SMP BUANA Waru Sidoarjo has high reliability, it is looked from the index of difficulty of the test.

Based on those results, we can generally conclude that index of difficulty in SMP BUANA Waru are good with index of difficulty value 0.30 – 0.70. Although the result in 7B shown that most of items are categorized as easy item test with 62%. Those items may be used again for the next test. However, there are so many numbers of items test that shown as easy test but there are some difficult items to. Both the difficulty and the easy one must be revised because they cannot work as expected to measure students' progress and which students that able to catch the material that has been taught by their teacher well. But the difficulty item test can make the student confuse and spend their time only focus to answer certain items and make them become lazy because they think that they have studied hard but they still cannot do the test well, so their motivation to study decrease. In the other hand, the easy item test do not give chance for students to increase their thought to think hard, so they will be lazy to study. Whereas, index of discrimination of English Multiple

