ITEM ANALYSIS OF ECONOMIC SEMESTER TEST OF GRADE XI IN ALL PUBLIC AND ISLAMIC SENIOR HIGH SCHOOLS IN BUKITTINGGI

Nita Sofia
Universitas Putra Indonesia (UPI) YPTK Padang
nitasofia111@yahoo.com

Abstrak

Analisis butir soal sangat penting dilakukan untuk memilih butir-butir soal yang sangat berkualitas. Tujuan dari penelitian ini adalah untuk mengetahui kualitas butir soal Ujian Akhir Semester (UAS) Ekonomi secara kualitatif dan kuantitatif. Penelitian ini menggunakan pendekatan deskriptif. Sampel penelitian ini adalah 256 lembar jawaban siswa yang mengikuti UAS Kelas XI di SMA/MA non RSBI se-Kota Bukittinggi. Teknik analisis yang digunakan adalah analisis kualitatif dan kuantitatif. Analisis kualitatif berupa analisis berdasarkan aspek materi, kontruksi dan bahasa dengan menggunakan lembar telaah. Analisis kuantitatif berupa analisis berdasarkan reliabilitas tes, tingkat kesukaran, daya pembeda, dan fungsi disktraktor dengan menggunakan program Anates versi 4.00. Hasil penelitian menunjukkan bahwa (1) dari analisis kualitatif dapat diketahui butir soal telah disusun secara baik dari aspek materi, kontruksi, dan bahasa sebanyak 37,5%; (2) berdasarkan analisis kuantitatif diperoleh koefisien reliabilitas butir soal yang tergolong rendah yaitu sebesar 0,67; tingkat kesukaran butir soal yang mudah sebanyak 30%, sedang sebanyak 57,5%, dan sukar sebanyak 12,5%, hanya sebanyak 47,5% daya pembeda butir soal yang berfungsi dengan baik; distraktor telah berfungsi dengan baik sebanyak 45% .

Kata kunci : analisis butir soal, analisis kualitatif, analisis kuantitatif.

Abstract

Item analysis is very important to select item which have good quality. The purpose of this research was to analyzed the quality of economics final exam using qualitative and quantitative approach. This study was descriptive research. It conducted in all senior high schools non pioneering international standard high schools in Bukittinggi. The data in this study were 256 student answer sheet in economic final exam for XI grade. Data were analyzed with qualitative and quantitative technique. The qualitative analyzed using the theoritical item characteristic based on material aspect, contruction aspect, and language aspect. The quantitative analyzed using the theoritical item characteristic based on reliability test, difficult item, item discrimination, and functioning distractor using anates version 4.00 program. The result show that (1) The item characteristic using the qualitative reviews are 37,5% tes in good category to material, construction, and languange aspect; (2) Based on quantitative analyszed coefisien of reliability that classified into low part is 0,67; the level of difficulties of easy questions as much as 30%, standard is 57,5, and difficult is 12,5%; only 47,5% item discrimination of question that well functioning; distractor are functioning properly as much as 45%.

Keyword : Item analysis, qualitative analysis, quantitative analysis.

INTRODUCTION

The learning process at school involves teachers and students. Both components work together harmoniously to achieve educational goals. Teachers as professionals are obliged to plan a focused and orderly learning process so that students can learn well, to implement learning based on a set of specific rules that help guide learning activities, and to assess and evaluate learning outcomes.

Assessing is the process of collecting and processing information to determine student learning outcomes. Nana (2008: 20) says "Assessment in education at least includes assessment of the program or curriculum, assessment of teaching and learning, and assessment of learning outcomes." One instrument for measuring frequently used learning outcomes is a test. According to Anas (2009: 67) "A test is a way or procedure in measuring and assessing of educational field in the form of assignment or a

series of tasks involving questions or orders by the participant so that the value generated can symbolize the behavior or achievement of the participant. "There are several characteristics that must be possessed by the test instruments of learning outcome; they must be valid, reliable, objective, and practical (Anas, 2009: 93).

One of most commonly tests used by teacher to assess and evaluate students' learning outcome is final semester test. The semester test of Economic subject in all Public and Islamic Senior High Schools in Bukitinggi is the same, except for pioneers of International-based Senior High School (RSBI). The test maker is appointed from Economic teacher of one of senior high schools in Bukittingi every semester. The semester test of Economic subject in July-December 2011 period was constructed by an Economic teacher of Islamic Senior High School (MAN) 2 Bukitinggi.

According to the teacher, the test was made by herself without any assistance from Economic teachers of other schools. They have fully handed it over the appointed teacher. It means the test was made only by one person. This teacher only considered her students' abilities and the lesson she taught in that school, so it might not be suitable for students of other schools.

Based on the data obtained from each school, students who have passed the minimum standard criteria were only 16% of all students in all public and Islamic senior high schools, except RSBI schools in Bukittinggi. It seemed that students got difficulties in answering the test items. It was assumed that there were some problems such as the items were not suitable with the lesson, the language was not appropriately used, the level difficulty of test was high, the discrimination index was inaccurate and the distracter efficiency did not work properly.

According to the interview results of all Economic teachers who teach at grade XI, the semester tests have never been tried out to see their validity and quality. Test item analysis is necessary to create a good quality of test. It is in line with Asmawi (2001: 172) who states some reasons why test item analysis is needed: to determine the strength and weakness of the test items, so the good ones can be selected, to provide information about test items' specification to make it easier to set the test, to determine the problems talked in the test items, to be a tool to assess the test items in the test collection. The test items can be analyzed qualitatively and quantitatively (Eko, 2009:94). The test items are analyzed qualitatively by considering material, language and construction; meanwhile they can be analyzed quantitatively by using empirical data to see the level of reliability, the level of difficulty, discrimination index, and distracter efficiency of each item.

If the test items were not analyzed before they are given to the participant, the quality of the test would be obscure to be used as the assessing tool, so the result could not reflex the students' ability. A good instrument is very important to grant the qualified evaluation for students. Good test items should have good criteria of materials, language, and constructions. Moreover, a good test item should have reliability index bigger than 0.7, level of difficulty as 0.3 to 0.7 (Anas, 2009,372), discrimination index as 0.3 (Djemari, 2008:143) and good distracter efficiency in which they are chosen by at least 5% of all participants (Suharsimi, 2009:220).

Therefore, by doing test item analysis it is expected that the quality of Economic semester test of of grade XI in all Public and Islamic Senior High Schools in Bukittinggi can be analyzed qualitatively and quantitatively.

RESEARCH METHOD

This research was descriptive research. The data were obtained from evaluation sheets from evaluators, final semester tests, answer key, students' answer sheets, students' score, and syllabus. The technique of data collection was documentation. The data were analyzed qualitatively and quantitatively. Qualitative analysis focused on material, language and construction of the test, while quantitative analysis concerned on the level of reliability, the level of difficulty, discrimination index, and distracter efficiency.

RESULT AND DISCUSSION

There were different results found from qualitative analysis. According to materials, it was revealed that there were 77.5% of test items that have fulfilled the requirement of good test. In other word, more than a half of test items were classified good. From the four criteria, the lowest percentage was on the test items which were not suitable with the indicators. 20% of all test items was not in line with the materials taught by teachers. As Djunaidi (2008) stated that 'materials aspect relates to the science of what is questioned as well as stratified level of thinking'.

Based on construction aspect, only 45% of all items can be considered as good test items. The writing technique in constructing the test had the lowest percentage; that was 17.4%. According to Djemari (2008:137), 'construction aspect relates to the technique of writing the test'. Most items were written in form of number, or were not organized in correct order and chronology. Another weakness was the unsuitability between the logical thinking on each item and that in the basic competences and indicators. In other word, the items tested were different with what teacher taught. As Suharsimi (2009:67) stated that a good construction of a test item can measure logical thinking as stated in Specific Instructional Goal. Moreover, the result of language analysis showed that some items of the

test did not follow the proper structure of Bahasa Indonesia.

From all qualitative aspect analyses toward the items of Economic semester test of grade XI, it can be revealed that 62.5% of items did not fulfill the criteria of good test. Among the three aspects analyzed, construction had the lowest level in which the items cannot be used properly to assess logical thinking of students.

Furthermore, the quantitative analysis was conducted to see the reliability, level of difficulty, discrimination index and distracter efficiency. The analysis of students' answer sheets showed that the reliability coefficient of the items was 0.67. In other word, the test items were not accurate to assess students' ability because it has low reliability index which it should be more than 0.7. Djemari (2008:31) assumed that the instruments used are suitable, but the interpretation of the correct answer can be different due to the way test conducted, emotional factors, or subjectivity of the test. Students' ability which is not homogeny can also lead to the low reliability index. In line with Grounlund in Dimiyati and Mudjiono (2006:196) that objectivity can affect reliability. In this case, students who join Economic semester test are from different schools which have different capability of students. If the school has more intelligence students, the test will have high result.

The level of difficulty analysis of this test item showed that 57.3%of items had moderate level, not too difficult or too easy. The coefficient level of difficulty was 0.30 to 0.7. Suharsimi (2009:207) states 'a good test is not too difficult nor is too easy'. An easy item cannot stimulate students to work hard to solve it, while a difficult item can lead to reduce students' motivation to solve it.

Another analysis to conduct was analyzing the discrimination index to determine students who have low ability and those who have high ability. The result showed that 47.5% of test items had good discrimination index. It meant that the items cannot accurately determine the ability of the students. As Ngalim (2009:121) states, 'discrimination index is used to determine students of upper group and those of lower group'.

The last analysis was the efficiency of distracters provided in Economic semester test items. It showed that 45% of items had good distracters which meant they were chosen by at least 5% of the test participant. Anas (2009:135) argued one weakness of multiple choice tests; that is the participant can speculate or guess the answer. The options provided can make students able to predict the answer when they do not know the exact answer. If their guessing are correct, it means the test is not accurate in assessing their ability.

It can be tailored from all analysis result that only 10% of test items can be used and fulfill the criteria to evaluate students' learning outcome. 67.5% of items needed to be revised and 22.5% of them were not appropriate to use. This condition occurred because only one Economic teacher constructed the test. It is strongly suggested that the schools have a team of each school to construct the test, so that the team can verify the items whether they are suitable to be use in each school. Furthermore, the items have never bee tried out, so the weakness of the test was unknown and not revised before it was given to the students. It is supported by Eko (2009) who states 'before the test is given to the students, it is necessary to conduct try out to improve the quality of the test item'. This can happen when the teachers did not know the step to construct and develop the test. As stated by Djemari (2008:88), there are 9 steps to construct test: arranging the test specifications, writing the items, analyzing the items, trying out the items, analyzing the items, revising the items, constructing the test, conducting the test, and analyzing the test result'. If these steps are not fulfilled yet, the test cannot be used to evaluate and assess the students learning outcome.

CONCLUSION

Qualitative item analysis of Economic semester test of grade XI in all public and Islamic senior high school in Bukittinggi showed that 37.5% of items have been well-constructed seen from material, construction, and language aspects. Quantitative analysis revealed that (1) the reliability coefficient was 0.67 which was lower than the good standard was, 0.7; (2) the level of difficulty showed that 30% was easy, 57.5% was moderate, and 12.5% was difficult; (3) 47.5% of test items had good discrimination index; (4) 45% of test items had good distracter efficiency.

Qualitative item analysis of Economic semester test of grade XI in all public and Islamic senior high school in Bukittinggi showed that 37.5% of items have been well-constructed seen from material, construction, and language aspects. Quantitative analysis revealed that (1) the reliability coefficient was 0.67 which was lower than the good standard was, 0.7; (2) the level of difficulty showed that 30% was easy, 57.5% was moderate, and 12.5% was difficult; (3) 47.5% of test items had good discrimination index; (4) 45% of test items had good distracter efficiency.

REFERENCE

Anas Sudijono (2009). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Raja Grafindo.

Asmawi Zainul (2001). *Penilaian Hasil Belajar*. Jakarta: PAU-PPAI-UT.

Djemari Mardapi (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Jogjakarta: Mitra Cendikia.

Djunaidi. 2008. http://evaluasipendidikan.blogspot.com/2008

/06/analisis-butir-soal.html. 15 Januari 2012.

Eko Putro Widoyoko (2009). *Evaluasi Program Pembelajaran*. Yogyakarta: Pustaka Belajar.

Murnialis (2003). *Teknik Evaluasi Hasil Belajar*. Jakarta: Grafindo Persada.

Nana Sudjana (2008). *Penilaian Hasil Proses Belajar Mengajar*. Jakarta: Rajawali Pers.

Ngalim Purwanto (2009). *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya

Suharsimi Arikunto (2009). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Rineka Cipta.

AUTHOR'S HISTORY

Nita Sofia, S.Pd., M.Pd.E. Nita lahir di Payakumbuh 23 September 1989. Beliau merupakan lulusan sarjana pendidikan ekonomi Universitas Negeri Padang pada tahun 2012. Tahun 2013 beliau memperoleh beasiswa program pendidikan calon akademi komunitas (PPCPAK) dari DIKTI selama satu tahun. Selama program tersebut beliau ikut serta mengajar di Politeknik Negeri Padang dan Akademi Komunitas Dharmasraya. Tahun 2015 beliau menyelesaikan magister pendidikan ekonomi dengan jurusan pendidikan akuntansi di Universitas Negeri Padang. Dan pada tahun itu juga beliau mulai menjadi dosen di UPI YPTK Padang pada jurusan Akuntansi.