

Teachers' Ability in Designing Test Assessments

Tatan Zenal Mutakin^{1*}, Arif Rahman Hakim²

¹*Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indraprasta PGRI, Jakarta, Indonesia*

²*Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indraprasta PGRI, Jakarta, Indonesia*

*Corresponding author: zmtatan74@gmail.com

ABSTRACT

The research is aimed at finding out information about the teachers' ability to design test assessments in class. The research approach applied a qualitative with descriptive analysis method. The research was conducted in three cities, Jakarta, Bekasi, and South Tangerang, with sample of 31 elementary school teachers. The results of this research were 66.4% of teachers knew the aspects of the content domain and cognitive domain in the curriculum, 78.49% of teachers designed and trialed the questions to be tested. However, 64.52% of teachers did not implement the cognitive domain analysis to questions to be tested. Those cases contributed to impact on the quality of low-test questions where 88.45% were categorized as low order thinking, 11.55% were categorized as middle order thinking, and 0% categorized as high order thinking. In that way, the students became unfamiliar to answer high order thinking questions. The research implied that it is necessary to evaluate the questions designed by the teachers to achieve high competitiveness. The research recommended teachers to upgrade the ability on how to design high-competitiveness test questions.

Keywords: *High competitiveness, Teachers' ability, Test assessments*

1. INTRODUCTION

A learning process can't be separated from input, process and output. The input of the learning process consists of students and teachers. The learning process is a method or media that implemented by the teacher to explain the material to students based on the curriculum applied. While, learning output is the result of the learning process. The learning outcomes are indicators of the learning process whether it is appropriate or not for the curriculum goals. That is why a final evaluation of the learning is needed.

Learning evaluation is very important to measure the development of teachers' quality in a country, especially in Indonesia. This is stated in UU SISDIKNAS number 20 of 2003, chapter XVI about evaluation, accreditation and certification in article 57 paragraph 1, "Evaluation is carried out in the context of national teacher's quality control as accountability of teacher organizers to interested parties. Arifin [1] stated that the purpose of evaluation is as a selection event, placement, diagnosis and remediation, feedback, interpretation of the norm reference, benchmark reference, motivation and tutor improvement of programs and curriculum, formative evaluation a summative and theory development.

Evaluation goals according to Sudijono [2] is to obtain evidence data that will be a guide to know the level of ability and student's success in achieving learning goals after taking the learning process. Furthermore, Khaerudin [3] mentioned that there are 2 learning evaluation, namely: 1) establish teaching competencies, and 2) improve the learning process. While in the learning outcomes, evaluation have several aims: 1) identify the difference in students' abilities, and 2) measure students' abilities individually or in group.

In some literature, evaluation has a broader scope than assessment. Suryanto [4] said that learning evaluation is a broad assessment, that is all components in learning program, while assessing individual scope is named assessment. Assessment has a definition on activity to gather information of the student learning outcome obtained from various types of bills and processing information to assess learning outcomes and development of student learning. Assessment process can't be separated from measurement process with a measuring instrument. According to Arifin [1] instrument has important functions to find out the effectiveness of the learning process. Meanwhile, Mulyadi [5] stated that measurement process includes 2 things namely measurement and test. When conducting an assessment, teacher must take measurement by using a tool called a test. Figure 1 displays the relationship among them.

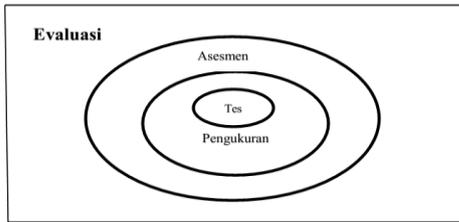


Figure 1. The relationship of evaluation, assessment, measurement, and test.

Hamid Hasan in Arifin [1] explained that test is a tool to collect the data which designed specifically. The implementation of the test must be properly planned and designed following the specific steps. Sudaryono [6] explained that steps for designing the test are: 1) set test objectives, 2) curriculum analysis, 3) textbook analysis, 4) make a grid, 5) writing specific instructional goals (basic competencies), 6) write questions, 7) review the question, 8) limit the test reproduction, 9) test run, 10) analysis of trial results, 11) revise questions, and 12) finalize the test questions.

Judging by the way it was made, test was distinguished into two types: 1) standard test, and 2) nonstandard test (made by the teacher). Standard test is a test which designed by experts carefully and includes academic goals which examines the level of validity and reliability. Meanwhile, nonstandard test is a test made by the teacher for its own scope. In general, formative and summative test which is made by the teacher have not been analyzed by the aspects of validity and reliability [7]. Based on history, standard test was applied in Indonesia since 1965-1979, known as the state exam; then in 1980-2000, state exam was replaced by final national evaluation of learning (EBTANAS); then in 2001-2004, it became national final exam (UNAS). In 2005-2010, the graduation score had a minimum target; then in 2010 until now, it has been known as national exam (UN). In 2002, EBTANAS in Elementary was abolished according to the national education minister decree number, 11/u/2002. The deletion was based on consideration that 9 years compulsory education program did not need EBTANAS. Since EBTANAS removal, one of the teachers' competencies was needed that was ability to design and develop question to be used in the final school exam (UAS). To produce quality test, special abilities of the teacher are needed in writing questions, besides the ability to develop questions continuously [8].

Standard test that is applied in our country today is the national exam (UN) for grade 9 and grade 12. In the past few years, the UN has been developed into a computer-based national exam (UNBK). Besides the UN or the UNBK, assessment using standard test is also applied for several Indonesian students through participation in an assessment survey conducted by Program for International Student Assessment (PISA). The PISA is referred for 15 years old student, and also Trend in International Mathematics and Science Study (TIMSS) that is designated for grade 4 and grade 8 student. In 2015, a survey result on Trend in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) ranked

Indonesia at 45th level out of 50 TIMSS survey participant countries; at 62nd for science and 63rd for mathematics from 69 survey participant countries.

As we know that assessment implementation using the standard test that is administered by the government only applied in the final semester of junior high school and senior high school, it means that assessment in the elementary level and early semester at junior high school that is known as a school exam, is using a test instrument that is designed by the teacher. Widodo [9] said that preparation of school exam test aims to measure students' success in mastering basic competencies that can be treated as indicators to find out the students' weaknesses. However, a good test instrument must fulfill several requirements that are quite difficult, require knowledge, skills and high accuracy.

2. METHOD

The population in this study were the teachers in surrounding Jakarta that represented by 31 teachers of elementary level in Jakarta, Bekasi and South Tangerang. The sample was selected by using the random sampling. The research instruments covered a questionnaire consisted of 24 questions that included 3 aspects, namely: knowledge, implementation, and feedback; fill in data consisted of 10 steps on test design; and school's documents with 13 examples of mathematics problems for grade 4. The research applied a qualitative approach with descriptive analysis method. The qualitative approach aimed at looking for a social phenomenon in other education sciences in deeper analysis. The qualitative approach was directed to produce improvement of work quality and benefits in academic purposes [10].

3. RESULTS AND DISCUSSION

3.1. Research Result

This study employed 3 data sources: questionnaire, fill in data and school documents. The questionnaire data consisted of 3 aspects, namely: aspects of knowledge, aspects of implementation, and aspects of feedback. Fill in data consisted of 10 related questions and steps for designing the test. For school documents, they consisted of example of mathematics problems for grade 4 in the odd semester of 2018/2019 school year.

Due to the data sources, the results would be described in three different sources. The results of questionnaire indicated that 65.35 % of teachers knew the technique to design the test questions, 73.84% of teachers carried out steps for designing a test, 64.52% of teachers carried out feedback after doing the test. The detail results were displayed in Table 1.

Table 1. Teacher's ability to design test question

No	Components assessed	Score (%)
1	Aspects of knowledge	65.35
2	Aspects of implementation	73.84

3	Aspects of feedback	64.52
4	Questions are taken from the textbook or students' worksheet or questions' bank	58.06
5	Less understanding the cognitive domain of questions	60.75

Based on the Table 1, several results showed that it needed to be further studied from the results of the questionnaire. The first, 58.06% of teachers didn't design question but took them from the textbook or the students' worksheet and question bank. Second, 60.75% of teachers had less understanding to the cognitive domain of the designed questions.

Continue to the results of fill data that indicated 67.47 % of teachers didn't do cognitive domain analysis to questions that he/she designated, and 67.74 % of teachers didn't revise the results of the trial analysis, as seen in Table 2.

Table 2. Cognitive Domain analysis and The Trial Results

No	Component assessed	Score (%)
1	Didn't do cognitive domain analysis	67.74
2	Didn't revise results of the trial	67.74

The final result was the analysis of example questions made by teachers on school year 2018/2019 in odd final exam. The result showed that the domain aspects of all questions that designed by the teachers had been according to the material required by the curriculum. Meanwhile, cognitive domain aspects showed that the quality of the questions designed by the teachers was still low with the details of 88.45 % categorized as low order thinking, 11.55 % categorized as medium order thinking, and there were no questions categorized as high order thinking. The results of this analysis were directly proportional to the result of the study of point B that obtained 67.47% of teachers didn't do cognitive domain analysis to questions that he /she designed. Table 3 displays the results.

Table 3. Cognitive analysis test questions

No	Cognitive domain test	Score (%)
1	Low order thinking	88.45
2	High order thinking	11.55

Source: 2019

3.2. Discussion

Assessment basically aimed to get information of the process and the learning outcomes from students and teachers during the learning process both in inside and outside the classroom. The results of the learning process were the master indicators for basic competition which

had been set in the curriculum. To measure the results of the learning process, a measuring instrument was needed that called a test instrument. A qualified test instrument would provide precise information about the students' abilities for competencies that had or had not been mastered.

Based on the way of designing test, there were two types of test: the standard test and the non-standard test that developed by the teacher. The standard test was a test designed by the experts, that was particularly designed by the government through the national examination in primary and secondary education. The standard test that is applied in our country known as National Examination for grade 9 and grade 12. Though, in the last few years, the national examination has developed into the computer-based of national examination (UNBK). Meanwhile, in the elementary level, school final exam (UAS) has been applied since 2002. For the nonstandard test or test designed by teachers, it often administered in formative and summative test for elementary, junior high school and senior high school.

Teachers' ability to compile test instrument is very important since the quality of the test instruments can measure accurately the students' ability in mastery learning materials in accordance with the competency indicators. Khaerudin [3] in a study said that a way to improve the most effective learning process is evaluating the test of learning outcomes based on the learning process itself. A good learning outcome is if it has met specific criteria and have high reliability and sustainability. A test that has high reliability is the test that produces measurement with steady and stable result. Meanwhile, the continuous means that assessment is conducted continuously to obtain an overview of students' competency achievement in a certain period.

One thing that needs to be considered in making a test, it must refer to the determination of basic competencies, contains aspects of content and cognition domain. Besides, the test must be done with good planning (by design) by following steps that have been determined. Sudaryono [6] explained that steps of design a test as an evaluation tool are: 1) set test, 2) curriculum analysis, 3) textbook analysis, 4) design grids, 5) writing specific instructional goals (basic competencies), 6) question writing, 7) review the question, 8) produce limited test, 9) trial test, 10) analysis of trial results, 11) revise the question, and 12) finalize the test questions.

The results of the study showed that the tests were designed by elementary teachers in Jakarta, Bekasi, South of Tangerang. The tests were still dominated by mastery of the content domain, meanwhile the cognitive domain was less noticed. Research data stated that 67.74 % of teachers didn't do cognitive domain analysis to questions that they designed and 67.74 % of teachers didn't revise questions of trial analysis. That had an impact on the question's quality. They were still low in details where 88.45 % categorized as low order thinking, 11.55 % categorized as medium order thinking and there was no question in high order thinking.

The above results reinforced the previous research conducted by Widodo [9], that said only 42% of teachers had the ability to design questions for final test well and

the others took questions from question bank or internet. It was also supported by Fatmayani's research [11] that stated three types of the teachers' ability in designing the test. The first, the level of teachers' ability to arrange very high question (24%), medium (28%), low (25%), very low (1%). Second, teachers' ability to develop low cognitive level was good: C1 (29.5%), C2 (23.8%), C3 (13.6%), though, the high cognitive level was very lack, C4 (1.14%), C5 (2.27%), C6(9.09%). The third, the teachers' ability to analyze question was still low where the level of items difficulty (10.23%), different items (3.41%), effectiveness of tricky (10.23%), question validity test (11.36%), and reliability of test questions (4.54%).

Refers to aforementioned research results, it is very reasonable that the mathematical abilities of our students still left behind compared to the students in other countries. Our test material was still struggling with the content domain, meanwhile the cognition had not been maximally applied. The assessment was conducted internationally has referred to high order thinking, while the cognitive ability of our students still low. In 2015, a survey result of Trend in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) ranked Indonesia at 45th level out of 50 TIMSS survey participant countries; at 62nd for science and 63rd for mathematics from 69 survey participant countries. The reason why our students' achievements are still left behind than other countries that is because our students are not equipped with high order thinking questions.

After EBTANAS has been removed since 2002, the assessment system for elementary level is entirely authorized by the school and teachers as organizer. It means assessment system and making question is entirely the task of teachers either in formative or summative test. This policy is a challenge for teachers to increase their capabilities continuously, especially in developing a good assessment based on the curriculum demands. A qualified assessment system will produce qualified evaluation and reproduce a right decision in establishing the learning process, the best one or not. Sudijono [2] said that learning evaluation goal is to obtain evidence data which will be a clue to know the level of students' ability in achieving learning goals after the learning process.

In general, strengthening steps are needed by the government to improve teachers' ability to review the curriculum as a whole, from the aspects of content domain and cognitive domain and implement it in the learning process or final evaluation system. By reviewing the curriculum and implementing it entirely, it is expected that learning goals as set and possibly achieved. The intended learning goals are an overview of the learning process and learning outcomes expected by students based on basic competencies. The abilities formulated in learning goals include abilities obtained by students during the learning process and the learning outcomes in a basic competency (Permendiknas, No. 41, 2007).

4. CONCLUSION

The ability in compiling test instrument is a competence that must be mastered by a teacher. A quality assessment of the test instrument can measure the accuracy of students' abilities based on the competency indicators set. Besides, the quality assessment of the test instrument is a way to improve effective learning process. A test instrument is good when it performs specified criteria, that are: 1) set test, 2) curriculum analysis, 3) textbook analysis, 4) design grids, 5) writing specific instructional goals (basic competencies), 6) question writing, 7) review the question, 8) produce limited test, 9) trial test, 10) analysis of trial results, and 11) revise the questions.

The research result was 66.4% of teachers knew aspects of content and cognitive domain in curriculum, 78.49% of teachers designed the test questions that would be assigned. However, 64.52% of teachers didn't apply cognitive domain analysis on questions that would be assigned. Those had impacts on the quality of low test' questions, 88.45% categorized as low order thinking, 11.55% categorized as medium order thinking and 0% categorized as high order thinking. The results contributed to the students' way of thinking that they were unfamiliar to answer high order thinking questions. Due to the results, it is necessary to evaluate questions designed by the teacher in order to achieve high competitiveness. The recommendation of this study is the teacher need to be provided by debrief in developing the test questions that are highly competitive.

REFERENCES

- [1] Z. Arifin, *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya, 2012.
- [2] A. Sudijono, *Pengantar Evaluasi Pendidikan*. Jakarta: RajaGrafindo, 2008.
- [3] Khaerudin, "Kualitas Instrumen Tes Hasil Belajar," *Madaniyah*, vol. 5, no. 2, 2015/8// 2015.
- [4] A. Suryanto, *Evaluasi Pembelajaran di SD*. Jakarta: Universitas Jakarta, 2016.
- [5] Mulyadi, *Evaluasi pendidikan: Pengembangan model evaluasi pendidikan agama di sekolah*. Malang: UIN-Maliki Press, 2010.
- [6] Sudaryono, *Pengantar Evaluasi Pendidikan*. Jakarta: Lentera Ilmu Cendikia, 2014.
- [7] A. Irwansyah, "Pengaruh Penguasaan Materi Matematika dan Pengetahuan Evaluasi terhadap Kualitas Tes Matematika Buatan Guru 2013," *Jurnal Inovasi Pendidikan MH Thamrin*, vol. 1, no. 1, pp. 50-61, 2017.
- [8] A. Purnomo, "Kemampuan guru dalam merancang tes berbentuk pilihan ganda pada mata pelajaran ips untuk ujian akhir sekolah (UAS)," *Lembaran Ilmu Kependidikan*, vol. 36, no. 1, 2007.

- [9] L. Widodo, "Upaya Meningkatkan Kompetensi Guru dalam Menyusun tes Hasil Belajar Melalui Workshop Penilaian di SMPN 2 Panti Jember," *JP3*, vol. 2, no. 2, 2012.
- [10] R. Indrawan and P. Yaniawati, *Metodologi penelitian kuantitatif, kualitatif, dan campuran untuk manajemen, pembangunan, dan pendidikan*. Bandung: PT Refika Aditama, 2014.
- [11] Fatmayanis, "Kemampuan Guru-guru Sejarah SMK dalam Menyusun dan Menganalisis Soal," *Manajer Pendidikan*, vol. 10, no. 5, 2016.