STUDY ON ABILITY OF ELEMENTARY SCHOOL TEACHER IN WRITING TEST ITEMS DURING THE PANDEMIC TIME OF COVID-19

Sri Wuryanti1 & Muhardis2
Center for Assessment and Learning, Ministry of Education and Culture
Jl. Gunung Sahari Raya no. 4, Jakarta, Indonesia.
E-mail: sriwuryanti03@gmail.com

Received: 24/10/2020; Revisions: 08/12/2020; Accepted: 22/12/2020
Published online: 22/01/2021

Abstract. The purpose of this study was to obtain information about the ability of elementary school teachers in composing test questions in private universities. This research uses qualitative research, which describes empirical reality using a descriptive approach. This study also uses a literature study, with documents to find some field data that occurs in the form of questions on several subjects. Meanwhile, the descriptive method is finding facts with the right interpretation, namely understanding the concept of the material, the ability to use the rules of the question. Meanwhile, the data analysis used was (1) the subject matter expert's judgment in interpreting some of the item test document findings, (2) theoretically and empirically the item had met the standard question rules. This research was found, during the COVID 19 pandemic, all parents or guardians have a 100% chance of assisting with distance learning as unprecedented. The sample consisted of nine subject teachers and the population was all fourth-grade teachers in primary schools as research subjects. So the assessment also uses offline and online so that it is found, (1) biased test questions, because the linguistic structure of the main sentence and explanatory sentences, is wrong in choosing questions with answers, (2) does not understand the rules about the questions: such as homogeneity of questioners, long and short answers, the use of standard Indonesian. Meanwhile, some of the test questions that were found to have problems were Indonesian, Natural Science, Character, and Boy Scouts. From 12 items, 4 items had material concept errors and 8 items error in the rules.

Keywords: test, understanding of concepts, rules of questions, subjects, assessment

INTRODUCTION

The COVID 19 pandemic requires parents to teach their children at home with distance learning. Learning takes place offline or online. Online learning methods include using google documents. Meanwhile, the offline method is to send learning materials such videos or Microsoft Power Point via Whatsapp. Likewise with the evaluation or assessment carried out by the teacher. Learning and assessment carried out by the teacher can be seen and understood by parents or pupils intensely. Parents try to transfer all the learning materials provided by the teacher to the children at home every day since the study from home takes place. However, the focus of attention and the problem of this paper is the content of the test items. These problems include: (1) some of the question items made by the teacher were inaccurate from the subject matter, and (2) several test items were not following the rules of writing the questions.

Regarded test items in a classroom learning assessment according to Setiadi (2010) states that the assessment process cannot be separated from the learning process, reflects the real world, uses several criteria or methods, and is holistic including aspects of learning (cognitive, affective, and sensory-motor).
Therefore, in determining test items in classroom assessment a teacher should not deviate from the concept of learning material, follow the rules of writing questions, and also increase reading literacy so that the test items can be used for the expected purpose. The objectives of the class assessment are (1) tracing: so that learning is according to plan, (2) checking: checking the weaknesses experienced by students, (3) searching: analyzing and reflecting which causes learning to be ineffective, and (4) conclusion: conclude whether students have mastered the competencies according to the curriculum.

A country that excels in the competition is not a country that simply has a wealth of natural resources. The country that will win the competition is a country that masters technology, a country that can educate its citizens to become thinkers and make them superior. The quality of human resources greatly affects the progress and prosperity of a nation. A nation with good and high quality human resources is expected to be able to compete with other nations in this era of globalization. Good quality human resources can only be done with a planned and systematic effort through quality education. Quality education is not as easy as turning your palms upside down. Learning in the classroom should leave conventional learning methods that merely pursue values and ignore competence. Teachers as educators must be equipped with various facilities, especially improving the competence of teacher in knowledge, teaching methods, as well as mastery of the curriculum, and mastery of assessment. The low reading interest and reading habits of students make reading skills low and that is what is now happening in our society today.

According to Suci (2020), students' reading literacy is the ability to access and retrieve information in simple text (150 to 200 words). They are able to determine the main sentence/main idea of a paragraph, determine the main character, and determine background evidence of the place/time/situation in the text. This is comparable to the 2018 PISA study, where the reading literacy skills of Indonesian students are at the level of detecting and handling conflict and representing information literally in the form of uncomplicated questions. Unfortunately, there are several abilities that students have not mastered based on the results of the two tests. From this research, it can be seen from the ability of Indonesian students to read literacy skills, whether it is related to teacher misconceptions in learning or assessment.

In the 21st century, known as the millennium era, information is widespread and communication becomes limitless. Millennials must be smart to filter the information needed. 21st century learning requires students to have essential skills, namely being able to think critically, creatively, and innovatively and solve problems, as well as collaborate and communicate effectively. These essential skills are needed by students as members of society and it is useful in life. Teachers still use many old patterns in teaching and conducting assessments, so that the ability of students to use reason to solve problems is still far from expectations. Related to this, a study is needed to determine the competence of elementary school teachers, so that input is obtained to improve the quality of education.

The problem raised in this study is how the ability of class teachers in compiling class assessments is seen from the suitability of the assessment with the subject matter, and the rules of writing questions.

The purpose of this study was to determine the test items made by the class teacher in compiling a class assessment by paying attention to the standard of the appropriateness of the assessment with the subject matter, and the rules of writing questions.

The rules for writing multiple-choice questions according to Nizam (2017) are (1) material: questions according to indicators, options are homogeneous and logic, and having the most correct answers. (2) construction: the subject matter is formulated clearly and firmly,
only the main formulation of the questions is needed, the subject matter should not give any clue to the correct answer, do not contain multiple negative statements, the length of the answer choices must be relatively the same, do not contain choices of all wrong or all correct in options, the choice of answers is based on the smallest or largest order, the figure or table must work, and the questions should not depend on the answer to the next or previous question. (3) Language: use the language according to the rules of the Indonesian language, do not use the locally applicable language, use communicative language, and do not repeat phrases that form one unit of meaning. And, an important note that does not offend SARA, does not have political content, pornography, violence, uses the name of an agency, and does not promote.

Meanwhile, the rules for writing essay questions according to the Puspendik team (2010) are (1) material: questions must be according to indicators, question boundaries must be clear, the content of the material is following the measurement instructions, and the content of the material is according to level. (2) Construction: sentence formulation using question words or commands, make clear question instructions, make scoring guidelines, and other things such as tables or figures must be presented clearly and meaningfully. (3) Language: uses simple and communicative language, does not offend students, and does not use words/sentences that cause multiple interpretations, uses good and correct language, takes into account language and culture, and does not use local languages. Puspendik also notes to immediately make scoring guidelines after writing the question sentences.

In fact, according to Deviana (2020), the condition of Indonesia is currently facing the 2019 coronavirus (COVID-19) pandemic. The importance of solutions is related to the role of teachers or educators to provide references such as social support to students in teaching and learning activities from home, this research conducted to provide research tools or instruments that can be used by educators or researchers by testing the validity of the social support construct measurement tool (Social Provision Scale) using the Rasch Model. The connection with this paper is how important it is for a measuring instrument to use a method that should be appropriate so that it will be meaningful to be used by students and teachers themselves.

According to Abdurrahman (2003), reading is essentially a complex activity that involves many things, not just pronouncing written text, but also including the use of understanding, visual activity, observation, and thinking. In line with the two experts, Zuchdi and Budiasih (2001) said that by reading a person can gain information, knowledge, and new experiences. To realize this, educators must be able to design and deliver learning material well, especially language skills, thinking, and mastery of subject matter. However, not all educators have this ability. According to the results of research conducted by Susilaningsih (2013), only 37% of educators can convey the subject matter clearly.

Meanwhile, according to the results of research conducted by Guritaningsih et al. (2000), the low quality of education is caused by low learning achievement, the difficulty of students in understanding and mastering subject matter, the inability to think critically and apply the knowledge obtained at school in different situations. The results of this research should be of concern to all parties, especially the Government, to improve the teaching quality of the educators. In classroom learning, educators must prepare themselves well in delivering subject matter to students. Educator readiness includes mastery of subject matter, programs, strategies, and learning methods. Learning strategies have a special role in the teaching and learning process in the classroom related to how much students can absorb subject matter, especially material related to critical and creative thinking processes or reasoning.

According to the Ministry of National Education (2010), there are basically three cognitive levels to get a specific picture of the ability level of students, namely: knowing and understanding, applying, and reasoning.
Students at the level of knowing and understanding have a minimum standard of ability to master lessons; students at the applying level have the ability to demonstrate knowledge and understanding of subject matter and can apply ideas and concepts in certain contexts; while students at the reasoning level have the ability to analyze, synthesize, evaluate, interpret, solve problems, and demonstrate original thoughts. To find out how much students absorb the subject matter properly, it can be done by evaluation or assessment. One of the assessment tools that can be used to measure the ability of students is a written test.

According to Cronbach (1960), Linn & Gronlund (1990), and Crocker and Algina (1986), testing is a systematic procedure for obtaining samples of student behavior. In the expert's opinion, it can be concluded that the test is a standard measuring tool used to obtain information about the abilities possessed by test takers. The test as a measuring tool must meet the prerequisites as a good measuring tool, namely the test must be valid and reliable. According to Brennan (2006), the validity of a test is the ability of a test to measure the characteristics or dimensions following what you want to measure, while the reliability or reliability of a test refers to the degree of a test that can measure various attributes consistently. A good test construction must meet these two requirements so that the test can provide a true picture of the condition of the testee (person) being tested, while according to Anastasi and Urbana (1988), test reliability is the stability of the score obtained by the same person when they are retested with the same test in different times or different sets of equivalent questions, or under different measurement conditions. The test as a measuring tool plays an important role in the teaching and learning process in the classroom. Test results can be used to determine student achievement and can be used to improve the quality of education. The test results can be in the form of grades or absorption (in percent) of the subject matter.

According to the Ministry of Education and Culture (2012), absorption is the percentage of students who can answer questions in the test, the absorption or mastery of this material is obtained from the calculation of the answers of students from each ability tested at each level. For the measured learning achievement test related to the cognitive level. Writing questions that require high reasoning requires good, creative mastery of subject matter, as well as expertise in writing questions, especially in problem construction.

The problem is that not all of the teachers can arrange good tests to be used as a reflection of student learning outcomes. As research conducted by Susiati (2019) states that SDN Yanti Jogoroto stated that 82% were not accompanied by grids and several grids were successful. Based on the analysis, if the questions are arranged from the wrong lattice it will result in the wrong questions. The wrong questions can certainly not be used to measure the expected achievement of basic competencies. Some of the main problems that cause the lack of grids that teachers have are due to a lack of understanding.

Meanwhile, research from Reality (2019) states that the difficulty of teachers in preparing assessment instruments is determining operational verbs, how to compile good language according to EYD (Enhanced Spelling) and communicative language editors, compiling aspects that are assessed ineffective instruments and psychomotor, adjusting between indicators and question instruments/aspects of assessment, and techniques for determining logical and homogeneous answer options (choices) on multiple-choice and matchmaking questions.

According to Leonardo's research (2014) based on the research team of Balitbangda (2005) of North Sumatra Province on the ability of teachers to make good evaluation tools for use in teaching and learning activities, it is known that 3.472% have very good abilities, 9.028% have good abilities, 59.722% have sufficient abilities good and 27.083% who have poor abilities. The ability of teachers to analyze the results of the evaluation to determine the abilities of students, it is known that 0% have very good abilities, 15.278% have good abilities, 52.778% have good enough abilities, and 31.944% of teachers who
have poor abilities. The teacher's problems in making test items are the wrong questions because there is no grid, difficulty in determining operational verbs, use of correct language, and only 3.50% make good evaluation tools, and only 15.3% can analyze the evaluation of student learning outcomes in the teaching and learning process.

**METHOD**

The population of this study was fourth-grade teachers at the primary school of the subject of the study and. the sample is subject teachers who have problems in the item test. The sampling technique was accidental sampling because the researcher happened to take a sample found in the field. Meanwhile, the research instrument was in the form of a daily test assessment manuscript in the form of secondary data. The data analysis technique used is descriptive qualitative, using content analysis by examining and reviewing the test items whether they are under the rules of writing standardized questions, by three subject matter experts from the Ministry of Education and Culture, namely Indonesian language experts, civic education and science.

**DISCUSSION**

The documentation that we collected for this research is that there are twelve test items from the subjects Indonesian, Science, Scouting, and Character. We compiled the test items as metadata, while in this discussion we wrote graphs and tables in the form of analysis of the test items found. The first discussion is the item test for Indonesian subjects, wherein these subjects there are some test items that experience problems, both the concept of the learning material or the rules of the problem, are as follows:

![Figure 1: Test items for Indonesian subjects](image)

Problems, namely: 2 with misconceptions in the material and 2 major errors in the rules of writing questions. While teachers in class subjects are taught by class teachers who are also subject teachers for grade four at the school. The errors that exist are about errors in the material concept, namely about Cannot interpret the main idea and explanatory idea and Not homogeneous. While the second error is about the rules for writing questions including Answer 1 2 3 is correct because there are no relationship between the answer choices and short or length of answer choices and repeating words.
Table 1. an overview of Indonesian questions

<table>
<thead>
<tr>
<th>No.</th>
<th>Subject</th>
<th>Concept</th>
<th>question writing rules</th>
<th>subjects and teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>Indonesian</td>
<td>Cannot interpret the main idea and explanatory idea</td>
<td>-</td>
<td>1 teacher</td>
</tr>
<tr>
<td>2)</td>
<td>Indonesian</td>
<td>Not homogeneous between main and explanatory ideas</td>
<td>Answer 1 2 3 is correct because there cannot be a relationship between the answer choices.</td>
<td></td>
</tr>
<tr>
<td>3)</td>
<td>Indonesian</td>
<td>-</td>
<td>Short length of answer choices and repeating words.</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Science class subjects are taught by class teachers, because science lessons are thematic lessons and classroom teachers who teach are the same as those teaching students in Indonesian.

While the second subject is science subjects for a general overview can be seen in the graphs and tables as follows:

![Figure 2: Test items for Science subjects](image_url)

From the two problems found to be problematic, there was 1 error in the material concept in terms of Cannot distinguish the middle ear with its subsections and another error in the rule of writing about questions about Repeating the word in the answer choices.

Table 2. an overview of Science questions

<table>
<thead>
<tr>
<th>No.</th>
<th>Subject</th>
<th>Concept</th>
<th>question writing rules</th>
<th>Subject and Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Science</td>
<td>-</td>
<td>Repeating the word in the answer choices</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>Science</td>
<td>Cannot distinguish the middle ear with its subsections</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Scout lessons are expected to provide character values to students, foster a spirit of hard work, togetherness and in this research. It is important to present it because scout teachers even though only two hours of teaching a week, are also required to have appropriate learning and assessment competencies students. Scouts are also taught to always have fun in accordance with scouting principles. Furthermore, the third is to analyze the scouting subject matter. These subjects in the 2013 curriculum are compulsory for all levels of education, including elementary schools. The graphs and tables can be seen as follows:
Scout teachers are from outside who have been certified as having abilities in the scouting field. While the problems that arise in making test items are not in accordance with the rules because they do not use standard language, good and correct in the request for answers.

<table>
<thead>
<tr>
<th>No.</th>
<th>Subject</th>
<th>Konsep</th>
<th>question writing rules</th>
<th>subjects and teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>4)</td>
<td>Scouting</td>
<td>-</td>
<td>Not using good and correct Indonesian</td>
<td>1 Scouting Teacher</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Next is the character learning subject taught by field teachers. The details can be seen as follows:

Character lessons are taught by one subject teacher. From the table, it can be described as follows: 1 is an error in the concept of the subject matter. While the other 5 mistakes were due to the use of violent language, inaccurate diction, and some that contained elements of discrimination for children with disabilities.
### Table 4. An overview of Charakter question

<table>
<thead>
<tr>
<th>No.</th>
<th>Subject</th>
<th>Concept</th>
<th>Question writing rules</th>
<th>Subjects and teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>5)</td>
<td>Character</td>
<td>-</td>
<td>The use of students’ body defect diction</td>
<td>1 Character Teacher</td>
</tr>
<tr>
<td>6)</td>
<td>Character</td>
<td>-</td>
<td>The answer choices are not homogeneous and the choice of cigarette diction</td>
<td></td>
</tr>
<tr>
<td>7)</td>
<td>Character</td>
<td>-</td>
<td>The short length of the answer choices and the writing is not good</td>
<td></td>
</tr>
<tr>
<td>8)</td>
<td>Character</td>
<td>-</td>
<td>Selection of drug diction</td>
<td></td>
</tr>
<tr>
<td>9)</td>
<td>Character</td>
<td>-</td>
<td>The selection of diction is not good in sentences</td>
<td></td>
</tr>
<tr>
<td>10)</td>
<td>Character</td>
<td>Non-specific questions</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Next are descriptions of all subjects. The details can be seen in the following graph and table.

![Figure 5. Test items for Indonesian, Science, Scout, Character subjects](image)

### Table 5. An overview of Indonesian, Science, Scout, Character question

<table>
<thead>
<tr>
<th>No.</th>
<th>Subject</th>
<th>Concept</th>
<th>Question writing rules</th>
<th>Subjects and teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>11)</td>
<td>Indonesian</td>
<td>Cannot interpret the main idea and explanatory idea</td>
<td>-</td>
<td>1 Indonesian Language Teacher</td>
</tr>
<tr>
<td>12)</td>
<td>Indonesian</td>
<td>Not homogeneous between main and explanatory ideas</td>
<td>Answer 1 2 3 is correct because there cannot be a relationship between the answer choices.</td>
<td></td>
</tr>
<tr>
<td>13)</td>
<td>Indonesian</td>
<td>-</td>
<td>Short length of answer choices and repeating words.</td>
<td></td>
</tr>
<tr>
<td>14)</td>
<td>Science</td>
<td>-</td>
<td>Repeating the word in the answer choices</td>
<td>1 Science Teacher</td>
</tr>
<tr>
<td>15)</td>
<td>Science</td>
<td>Cannot distinguish the middle ear with its subsections</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>16)</td>
<td>Scouting</td>
<td>-</td>
<td>Not using good and correct Indonesian</td>
<td>1 Scouting Teacher</td>
</tr>
<tr>
<td>17)</td>
<td>Character</td>
<td>-</td>
<td>The use of students’ body defect diction</td>
<td>1 Character Teacher</td>
</tr>
<tr>
<td>18)</td>
<td>Character</td>
<td>-</td>
<td>The answer choices are not homogeneous and the choice of cigarette diction</td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th></th>
<th>Character</th>
<th>Non-specific questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>19)</td>
<td>The short length of the answer choices and the writing is not good.</td>
<td></td>
</tr>
<tr>
<td>20)</td>
<td>Selection of drug diction</td>
<td></td>
</tr>
<tr>
<td>21)</td>
<td>The selection of diction is not good in sentences</td>
<td></td>
</tr>
<tr>
<td>22)</td>
<td>Total 4 8</td>
<td></td>
</tr>
</tbody>
</table>

The total sample consisting of 9 subject teachers, who made an error test item consisted of 3 subject teachers. Because, Indonesian and science subjects are taught by one class teacher because it is a thematic lesson.

From the total test items, material misconceptions are only 4 questions consisting of the interpretation of the main idea and explanation, as well as questions about sections or subsections in a material. And the remaining 9 questions are incorrect writing rules, including answer choices that still connect one answer choice with another, word repetition in the answer choices, the length and shortness of the answer choices are not the same, the use of diction is not the same. Correct, and contains elements of racial and religious beliefs as well as negative diction choices, and is found to be not homogeneous in the diversion in the choice of answers or homogeneity in the stimulation of questions.

From these data it is known that there are 4 subjects from Indonesian language subject teachers, namely Science, Scouting, and Character. Misunderstanding of the concept of subject matter where the mistakes include:

1) Cannot interpret the main idea and explanatory idea
2) Not homogeneous between the main idea and the explanatory idea
3) Cannot distinguish the middle ear from its subsections
4) Non-specific questions

Lastly, students will have no difficulty accepting new subjects. What is tragic is when the misunderstanding of subject concepts occurs in students at a lower level of education. Then, will get the wrong understanding on an ongoing basis. Whereas in the use of the rules for writing questions in general, several errors were found:

1) Answer 1 2 3 is correct because there are no relationship between the answer choices.
2) The short or length of answer choices and repeating words.
3) Not using good and correct Indonesian language
4) The selection of diction
5) The answer choices are not homogeneous and the choice of cigarette diction.

According to Arikunto (2013), the assessment aims to measure the ability of students, including the meaning for students is to know the extent of the ability to follow lessons, and the meaning for the teacher is knowing which students have the right to continue the lesson, right on target and knowing the maximum methods of teaching. And, the third is the meaning for schools to find out the conditions of student learning, the appropriateness of using the curriculum, and the assessment information used for guidance for subsequent years. If there is an error in some test items that do not follow the rules such as language use, improper diction, and use of sentences that are not the same length, this is generally very contrary to the purpose of the measurement itself.

In particular, some of these errors are (1) linking the one answer option with the next choice will confuse students (2) the short or length of the sentence in the answer choices will lead to the answer which is definitely correct and aesthetically less good, (2) the use of Indonesian
that is not standardized and making the wrong answer with the wrong Indonesian, will lead to several wrong perceptions in the meaning of the use of the language. Because this is an educational institution and students are always trained in good language, it would not be good to blame the use of language, (3) the use of some racism and poor diction in educational assessments must be avoided because this diction can be chosen in more languages. Meaning good, because learning and assessment teach good character.

CONCLUSION

The results of this study indicate that there are still misconceptions of subject matter and errors in the use of problem formulation rules. As mentioned above, if the test tool is not good then it cannot be used to measure learning outcomes maximally. Moreover, for learning reflection for students to observe students' weaknesses or strengths in mastering learning competencies or reflecting on the delivery of teacher learning. Therefore, the better practice is still needed to produce good and meaningful measurement test items.

Primary school teachers are classroom teachers or non-subject teachers, so on average, these teachers in schools must teach various compulsory subjects, such as; Indonesian, Mathematics, Social Sciences, Natural Sciences, and if Islamic schools are likely to teach some religious subject matter. Even though the subject matter is still basic, all need a good understanding. Therefore, classroom teachers should read a lot of basic material concepts and reproduce general literacy related to the subject matter carefully and thoroughly to enrich knowledge, as well as many discussions with forums between teachers in schools or groups of elementary school teachers so that the level of understanding in learning and assessment is more honed.

ACKNOWLEDGEMENT

We would like to thank Mr. Bagus Hary Prakoso who has inspired a lot for this article. Friends of researchers and engineers at Puspendik who cannot be mentioned one by one. Also for Mrs. Lilis, the head of the human resources division who is very supportive of the progress of Pusmenjar staff. Mrs. Asijanty, who is always open and frees our minds to encourage us all.

*****

REFERENCES


http://journal.uny.ac.id/index.php/jpep/issue/view/1308


http://journal.uny.ac.id/index.php/jpep/article/view/7173


