

CHAPTER I

INTRODUCTION

Background

In order to assess students' understanding and the efficacy of instructional practices, assessment is a crucial part of the teaching and learning process. At the conclusion of a learning session, such a semester or academic year, summative assessment is one of the most popular forms of evaluation to gauge students' progress (Brown, 2004). Summative evaluations are intended to give a clear image of students' subject-matter proficiency and are frequently used to gauge academic achievement and advancement. Due to its practicality in evaluating a large number of students, impartiality in scoring, and effectiveness in covering a wide range of topics, the multiple-choice question (MCQ) is one of the most commonly utilized formats in summative assessments (Nitko and Brookhart, 2011). Notwithstanding these benefits, the quality of MCQs' creation has a significant impact on how well they work as an assessment tool. Inaccurate judgments on students' learning results might result from poorly constructed multiple-choice questions that misrepresent students' true skills (Haladyna, Downing, and Rodriguez, 2002).

According to Nitko and Brookhart (2011), a well-structured multiple-choice question consists of three main components: the stem, the distractors and the correct answer (Haladyna et al., 2002). The stem is the main question or statement, which should be clear, concise and unambiguous. A good stem should avoid unnecessary complexity and negative wording that may confuse students rather than assess their true knowledge (McMillan, 2013). The distractors or incorrect answer choices should be plausible and based on common misconceptions or errors made by students. Distractors that are too obvious or nonsensical will not serve their purpose

as students can easily eliminate them, thus reducing the effectiveness of the question (Haladyna et al., 2002). Finally, the correct answer should be factually correct and not easily distinguishable based on word patterns or length differences. Studies have shown that poorly constructed MCQs often include answer choices that are disproportionately long or use specific key words, making it easier for students to guess the correct answer without really understanding the content.

In Indonesia, multiple-choice questions are a dominant format in national and school-based assessments, especially in primary schools. According to the Kementerian Pendidikan dan Kebudayaan (Kemendikbud, 2020), the use of MCQs is preferred due to their ease of administration and scoring, especially in large-scale standardised tests. However, research suggests that many MCQs used in Indonesian schools do not always meet quality standards, leading to concerns about their validity and reliability as an assessment tool. Yulianti & Suryani (2019) conducted an analysis of multiple-choice tests in Indonesian primary schools and found that a significant proportion of questions had ambiguous stems, ineffective distractors, and predictable answer patterns that allowed students to guess answers rather than demonstrate actual understanding. These deficiencies undermine the effectiveness of MCQs in accurately assessing students' knowledge and cognitive abilities.

Furthermore, the cognitive level of multiple-choice questions is another critical aspect to consider. Bloom's Taxonomy (2021), later revised by Anderson and Krathwohl (2001), cognitive processes are categorized into six levels: Remembering (C1), Understanding (C2), Applying (C3), Analysing (C4), Evaluating (C5) and Creating (C6). Research has shown that many multiple-choice questions in primary school assessments tend to focus heavily on lower order

thinking skills (LOTS), such as recall (C1) and comprehension (C2), rather than higher order thinking skills (HOTS), such as application, analysis or evaluation (C3-C5) (Brookhart, 2010). As a result, students may perform well on tests without necessarily developing deeper understanding or problem-solving skills. This overemphasis on rote memorisation rather than critical thinking runs counter to the goals of modern education, which aim to prepare students for more complex cognitive tasks (Anderson and Krathwohl, 2001).

An in-depth examination of the quality of multiple-choice questions is necessary because to their crucial role in basic education and the issues surrounding their creation. The purpose of this study is to investigate how multiple-choice questions are constructed for Grade 5 summative examinations at Agape Indah Elementary School. This school was chosen as a research location because it met several important criteria: it is relevant to the research focus, has good academic quality, is easily accessible, and is willing to cooperate with researchers. Selecting this location is expected to produce useful, applicable, and actionable findings.

To ascertain if these questions follow accepted guidelines for good assessment design, the study will specifically examine the suitability of answer choices, the efficacy of distractions, and the clarity of stems. In order to determine whether the questions accurately reflect students' learning at different cognitive levels and are in line with Bloom's Taxonomy, the study will also investigate the cognitive levels that are reflected in the questions.

By pointing out common flaws and offering suggestions for improving question quality, the study's findings should help elementary schools improve the design of their multiple-choice questions. Teachers can gain more precise insights

into students' comprehension and make better instructional judgments by making sure summative evaluations are designed well.

The actual conditions in this study were identified through a preliminary observation conducted prior to the main research. The preliminary observation was carried out by reviewing the documents of the Grade 5 English summative assessment at Agape Elementary School. The observation focused on the construction of multiple-choice questions, particularly the clarity of the stems, the quality of answer options and distractors, the alignment with learning indicators, and the cognitive levels of the items based on the Revised Bloom's Taxonomy. The results of the preliminary observation indicate that most of the multiple-choice questions were still constructed in a simple manner and did not fully follow the principles and standards of good item writing. Several items only contained brief instructions such as "*look at the picture*" without clearly stating the language competency being assessed. As a result, students tended to answer the questions based on visual guessing rather than actual English comprehension. In addition, the majority of the items measured only low-level cognitive skills, specifically remembering (C1), with limited variation in higher cognitive levels. These actual conditions show that the summative assessment questions have not functioned optimally as tools to comprehensively evaluate students' learning outcomes. Therefore, the findings of this preliminary observation serve as the empirical basis for conducting a more in-depth analysis of the construction of multiple-choice questions used in the Grade 5 English summative assessment at Agape Elementary School.

In addition, the questions used in summative assessment mostly only measure

low-level cognitive abilities, namely remembering, and do not provide a variety of thinking levels in accordance with Bloom's Revised Taxonomy. This condition shows that the assessment questions do not fully function as an evaluation tool capable of describing students' overall learning achievements. The author is interested in raising the title “An Analysis of Multiple-Choice Questions Construction in Summative Assessment for Grade 5 Students at Agape Elementary School” to systematically analyze the quality of the multiple-choice questions used and provide a realistic picture of the suitability of the questions with the principles and standards of good question construction.

Research Problem(s)

1. How do the multiple-choice questions in summative assessments for grade 5 at Agape Indah elementary school align with the principles of good questions construction?
2. Do the questions of multiple choice meet the standards of compiling good multiple-choice questions, such as clarity of the stem (main question), quality of distractors (distracting answers) and accuracy of the answer options?

Aim(s) of Study

1. This study aims to identify the multiple-choice questions in summative assessments for grade 5 at Agape Indah elementary school align with the principles of good questions construction
2. This study aims to identify wheter questions of multiple choice meet the standards of compiling good multiple-choice questions, such as clarity of the stem (main question), quality of distractors (distracting answers) and accuracy of the answer options

Significance of Study

The significance of the study is as follows:

Implication

This research contributes to the development of knowledge in the field of assessment, particularly in the construction of effective multiple-choice questions. The results of this study reinforce and apply previously developed theories, such as the item-writing principles of Haladyna, Downing, and Rodriguez (2002), which emphasize the importance of using clear language, selecting distractors that work, and compatibility with learning objectives.

In addition, this study also supports the use of the revised Bloom's Taxonomy (Anderson and Krathwohl, 2001) in designing items that measure not only factual knowledge but also higher order thinking skills. Thus, the results of this study can serve as a theoretical reference for further research in the area of question validity and the quality of assessment instruments.

Application

This study is highly relevant for application in learning activities, particularly in evaluating learning outcomes. In this context, the study's findings provide an overview of the quality of the multiple-choice questions used by fifth-grade teachers at SD Agape Indah in terms of the clarity of the stem (question formulation), the quality of the distractors (incorrect answers), and the accuracy of the answer key. Teachers can use these results to improve their methods for creating summative assessment questions and ensure that the questions accurately measure students' competency. According to Anderson and Krathwohl (2001), good assessment requires a deep understanding of the taxonomy of learning objectives and the appropriate evaluation techniques.

The study shows that analyzing multiple-choice questions is important not only theoretically, but also in daily educational practice. The research findings: help teachers improve their question-writing techniques, serve as a quality control tool for schools, and provide an academic foundation for the development of evaluation science. Implementing these findings is expected to improve the quality of assessments at the elementary school level, enabling them to objectively and fairly reflect students' learning achievements.