International Journal of Academic Multidisciplinary Research (IJAMR)
ISSN: 2643-9670
Vol. 10 Issue 7 July - 2026, Pages: 491-503

The Use, Misuse, and Abuse of Test Construction in Educational Assessment

Justin M. Cuyugan[1], Grace N. Alfonso [2], Erwin R. Dela Vega[3], Loiusse Wenchelle R. Ladrillo[4] Fatima Nicole R. Lejero[5], Reymond Q. Fajardo[6]

[1]Santo Tomas Campus, Pampanga State University, Philippines
justinmarfilcuyugan@gmail.com
[2]Santo Tomas Campus, Pampanga State University, Philippines
gracealfonso25@gmail.com
[3]Santo Tomas Campus, Pampanga State University, Philippines
erwindelavega17@gmail.com
[4]Santo Tomas Campus, Pampanga State University, Philippines
wenchelleladrillo@gmail.com
[5]Santo Tomas Campus, Pampanga State University, Philippines
lejerofatima@gmail.com
[6]Santo Tomas Campus, Pampanga State University, Philippines
rqfajardo@pampangastateu.edu.ph

Abstract: *Test construction is crucial in the endeavor of evaluating students' acquisition of knowledge. This prompted academics to explore the multifaceted nature of test construction. Notably, the examined papers highlight the necessity for ongoing research on test construction. Given these prevalent issues, the study investigates test construction practices used in science and mathematics assessments for grades four, five, and six over a period of four quarters to identify common problems as misuse and abuse of best practices. Based on indicators such as the clarity of item stems, readability, grammatical consistency, and the plausibility of distracters, the study sought to determine if there were any prevailing issues that could be classified into either misuse or abuse for both subjects. Through content analysis of the test papers, the study found that in science assessments, misuse was seen in readability (6 incidents) while item stem clarity had significant abuses (27 incidents), among others. Some misuses in math tests included items' level of readability (21 cases) and how alternatives were put together (4 instances), whereas some abuses included unclear item stems (57 cases) and having multiple correct alternatives (11 situations). These results emphasize the need for improved test construction methods aiming at enhancing validity and reliability of educational assessments. It concludes that this study identified many deficiencies notably with regard to poor wording or grammar in test items. This calls for better trained item writers; more exacting review processes; mechanization of portions of the test assembly process.*

Keywords— Test Construction, Common errors, Occurrences, Intervention.

1. INTRODUCTION

Poor test construction is like a sexual assault on the instruction. To commit violations of guidelines in constructing items is an exploitation of the quality of education. According to Zijlmans et al. (2019), test construction is a process that involves designing something from its inception all the way through to its use and scoring. Regarding the field of education. Creating a test requires a great deal of expertise and Monica S. (2021, March) noted that test construction plays an important role in education by demonstrating the strengths and weaknesses of the indicating cycle and assisting teachers in improving it. However, because of the significant increase in its use by educators, organizations, and legislators, test construction in educational assessment has become subject to new restrictions.

The relationship between abused and misused tests in educational assessment is a delicate issue that requires close examination. Assessing the degree to which learning objectives have been reached is one of the main duties of a classroom teacher or instructor, according to Adom, D. et al. (2020). The learner's competency in test construction cannot be disregarded to effectively measure the achieved instructional objectives. Even if the teacher is one aspect of test construction that can be abused and misused, creating assessments that truly measure students' comprehension and abilities still presents a number of difficulties around the world.

The study conducted at Silliman University in the Philippines, titled "The Conformity of Test Construction of the Achievement Test Papers of College Teachers: A Case Study" by dela Rama, P. A. (2022) indicates that not all teachers possess a high level of proficiency in adhering to test construction rules. It further highlights that the True or False type of questions is the most frequently utilized format. Nevertheless, certain teachers inadvertently reassess the caliber of their examination questions before administering

them. Additional concerns encompass the arrangement and directives that perplex students or are repetitive. Female teachers generally adhere to a greater number of test construction guidelines compared to male teachers, potentially due to intrinsic characteristics and differences in attitudes.

A study in Indonesia by Monica, S. (2021) concludes that the interview results suggest that there are certain practical issues wherein teachers have challenges in developing and conducting objective tests. The major issue is how to formulate the items for the objective test itself. The challenges included difficulty in picking which materials to assess, trouble in selecting plausible distractors for multiple-choice, difficulty in making students comprehend the test questions, difficulty in controlling the test duration, and difficulty.

In the same dimension, Inko-Taria et al. (2019) argue that the skill of educators, most especially with assessment, determines the quality of the exam developed, while Opie et al. (2021) assert that teachers' competence in assessing students' learning is frequently disregarded. However, beneath this apparently objective realm lies a tangled mass of intricacies within which the design, administration, and interpretation of tests can be possessed by individuals whose decisions may affect many others.

Science teachers had higher levels of self-assessment in test construction competencies for instructional decision-making, as compared to school administrators. Female science educators who possess higher degrees and have actively engaged in seminars on test creation are more experts in the areas of result administration, scoring, and interpretation. There is a positive correlation between the academic accomplishment of students in the field of science and the test construction competencies of teachers, Pastrana (2023). Moreover, Suchyadia et al. (2020) assert that teachers who are the intended beneficiaries of mentoring must not disregard their obligations in the field of education. In addition, researchers also consider the credentials of teachers to assess their ability to develop a test.

According to a study conducted in Nigeria by Opateye (2023) mentioned that teachers who do not hold recognized teaching credentials are not equipped with the skills needed to create high-quality materials that meet the essential standards for validity and reliability. Over time, those instructors were taught most test-related topics, including validity, reliability, and item analysis. However, teachers with fewer years of experience were not afforded the same opportunity. Furthermore, elements like the kind of test that teachers utilize to create educational assessments may also be taken into consideration as problematic.

Assessment components are internally coherent when they are mutually supportive. This means that formative and summative tests should be related to each other. Hence, the efficacy of evaluation would be constrained by the characteristics of the broader framework in which it is integrated, specifically, the substance, structure, and arrangement. Therefore, the efficacy of the assessment would be constrained by the characteristics of the broader system in which it is integrated, including the content, format, and design (Areekkuzhiyil, 2021). KLEU (2024) emphasized that validity and reliability are crucial aspects of any assessment. Essentially, it implies that assessment should be both precise and accurate. Furthermore, the efficacy of test construction remains a persistent concern in various regions of the globe.

Improving teacher assessment literacy is crucial in Georgia. Georgia has made efforts to enhance its utilization of assessment to enhance student learning, but the results have been variable. This report proposes a range of measures to achieve further advancements in this domain. Countries, including Georgia, have encountered difficulties such as a lack of alignment between the curriculum and assessment methods, as well as insufficient assistance for educators striving to implement reforms in their classrooms. This was asserted by R. Li, H. Kitchen, B. George, M. Richardson, and E. Fordham (2019). Moreover, teachers select specific types of tests during the construction process to ensure that they are successful and connected with their intended goals.

According to Kumar et al. (2021), one type of test is the multiple-choice questions (MCQs) which are widely employed in college and higher education medical assessments because they are efficient, reliable, and easily standardized. However, this is the most frequently devised, utilized, and occasionally abused by classroom teachers Opie et al. (2021). Furthermore, it is stated that a type of test is one of the factors in potentially misusing and abusing in educational assessment, but we have to consider the teachers' test construction skills as well in line with its practices.

A study by Quansah et al. (2019) investigates the teachers' test construction skills in some high schools in Ghana. The results revealed that most teachers had inadequate test construction skills. These findings have consequences for educational evaluation, as appropriate evaluation of learning outcomes is extremely important. Important given that test scores are utilized in decision-making. This was made clear when problems were discovered with the test's representativeness and relevancy of content, as well as the assessment tasks' fairness and dependability. According to Eni et al. (2019), test takers' failure on classroom tests isn't always the result of their tiresome nature, but rather of their teachers' poor test item creation. The quality of the test itself must satisfy a specific set of requirements in order for the measurements used in test item development to be deemed reliable. As a crucial aspect of test quality, educators are obliged to report on reliability on a regular basis. A test creator may wish to add or remove questions from an already-existing exam for a variety of reasons. Furthermore, a potential problem could be that the current exam is too short.

The interplay between teaching, learning, and assessment holds substantial importance in the field of education; it can be

likened to the participants in a three-legged race, as suggested by Butakor (2022). From this standpoint, Butakor (2022) underscores the importance and pertinence of classroom assessment methods, including those developed by instructors, so far as they facilitate student learning. The study's findings indicated that the mathematics test exhibited a satisfactory level of difficulty for one of its items, despite its low but acceptable reliability coefficient. The other item was both excessively difficult and simple.

Consistent with this, Ahyan et al. (2019) assert in their study "Students' Errors In Solving the PISA Mathematics Problem Using Newman's Error Categories" that difficulties with both transformational and process skills are common among students' errors. This is because students encounter challenges when attempting to convert practical or contextual situations into mathematical problems. In the absence of any pupil committing the encoding error. Moreover, there are still issues where educational assessments are being misused and potentially abused.

Multiple instances are provided in an article published in the prestigious journal Educational Researcher Rutkowski & Rutkowski that demonstrate the PISA project's omitted flaws and limitations, as well as the uncertainty surrounding their conclusions and recommendations. According to Sjöberg S. (2020), the primary factors considered inappropriate when assessing units were the absence of context, the inclusion of inappropriate context, cultural bias, reliance solely on school science, and the inclusion of deemed too advanced topics. Sjøberg, S. (2020) provides further clarification regarding the explicit indication that PISA does not evaluate academic knowledge or curriculum-related concerns.

Assessment influences directly what and how students learn and can either promote or inhibit effective learning (Bauer-Hoffbeck, 2019). The author argues that several publications that undertake critical analyses of the misuse and overuse of ILSA data would be valuable additions to this journal. This is due to the international impact that such studies have on the field of assessment and the opportunity they offer to gain a deeper understanding of students' learning. The researchers in the study "The Use and Abuse of Assessment, Assessment in Education" (2019) argue that extreme caution should be exercised when designing tests for educational purposes. While tests have the potential to enhance comprehension, their improper application can cloud perception, and their misuse can hinder learning. Moreover, the use, misuse, and abuse of test construction in educational assessment across the globe suggest possible issues that may arise in the Philippines as well.

The Philippines achieved the lowest ranking among 81 countries in the 2022 Programme for International Student Assessment (PISA), as stated in the report. Based on the December 2022 PISA results, the Philippines achieved an average score of 356 for science, 355 for mathematics, and 347 for reading, all of which ranked them third-to-last in their respective subjects. According to Sen. Sherwin Gatchalian, Benigno Aquino Ninoy High School in Makati City emerged as the sole public institution in the nation to achieve a passing score in PISA 2022, with 187 other schools failing the examination. Furthermore, factors derived from outcomes may vary because of various challenges and factors encountered by the Philippine education system and the quality of educational assessment, should any deficiencies exist.

In the Philippines, comparable issues persist but are aggravated by specific challenges inherent to the local pedagogical framework as well which is the same occurrence around the globe. A study about the Reshaping Assessment Practices in Philippine Teacher Education states that it is recommended that in reshaping test construction practices at this time, different contexts must be cogently considered so that reasonable changes would be better understood. As added by Cahapay, M. B. (2020), this is due to the improper application of assessments frequently occurring owing to insufficient resources dedicated to developing tests that accurately represent the culturally pertinent and linguistically suitable evaluation of students' knowledge.

This commonly leads to a dependency on imported standardized examinations that may fail to precisely reflect the experiences and learning of Filipino students. The pressure for students to perform well on high-stakes tests, such as the National Achievement Test (NAT) or Programme for International Student Assessment (PISA) as cited earlier for schools, can result in an instruction focused solely on preparing for the evaluation rather than fostering holistic scholarship experiences. Furthermore, this narrow focus risks neglecting important facets of education, such as socio-emotional growth and creativity, worsening inequalities in access to quality instruction.

The lack of appropriately tailored assessments fails to capture the diverse array of skills and understanding possessed by students of varied cultural and linguistic backgrounds. Barrientos (2023) states that teachers' understanding of objective test construction processes in teacher education programs was typically satisfactory. While teachers are well-versed in accurate test construction techniques for constructing the exam, they place a lower focus on preparing relevant test items. Nonetheless, reliance on examinations developed abroad neglects the local context and lived experiences of Filipino youth.

Conceptual Framework

The theoretical framework of the study is represented by a Venn diagram consisting of four circles in Figure 1. The core concept of educational assessment lies at the center of the diagram, which is also the focal point of this research. Around this central circle are three other circles whose key components are test construction; namely "Use," "Misuse," and "Abuse." While referring to appropriate and effective application of test construction techniques in educational assessment practices, the word "Use" appears within a circle. In contrast, above is another circle labeled "Misuse" showing instances where testing construction methods might be wrongly or ineffectively employed causing undesirable outcomes in assessment processes. Likewise, written above is another circle with the title 'Abuse', denoting extreme and damaging misapplication of test-construction methodologies that lead to biased or unfair assessments. This structure highlights distinctions between proper use, misuse, and abuse of testing methods for educational evaluation purposes. This framework acts as a compass for formulating strategies geared towards professional development.

This study investigated the use, misuse, and abuse of test construction in educational assessment from the perspective of constructivism, a theory developed by Jean Piaget. Constructivism highlights the active role of learners in constructing knowledge via their experiences and interactions. This research examined the alignment between teacher-made exams and constructivist ideas via analysis. It studied if the assessments promote students' use of existing knowledge and active involvement with the content in a significant manner, or if they just assess memory and neglect to capture the extent of students' comprehension. Exploring teacher-made tests from a constructivist perspective can provide useful insights on enhancing assessment methods to foster more profound learning and knowledge building.

The researchers aimed to explore the use, misuse, and abuse of test construction in educational assessment among teachers at the intermediate level of a certain elementary school.

Specifically, they sought answers to the following questions:

1.What is the use of test construction in line of education among elementary teachers?
2.What are the common errors elementary teachers make in test construction?
3.What ethical concerns arise from potential misuse or abuse of test construction practices?
4.What areas of professional development are most needed to enhance elementary teachers skills test construction

cope and Delimitation

The construction of tests in educational assessment plays a vital role in the education system. The quality of educational assessment, especially at elementary levels, is highly instrumental in students' academic growth and teachers' professionalism. The purpose of this study was to examine the use, misuse, and abuse of test construction in educational assessment within a specific school in Pampanga, to improve the implementation of assessment practices.

*Teachers.*The study aimed to discover areas for development by analyzing the test construction skills of intermediate-level teachers. This knowledge can be applied to design professional development programs that provide teachers with the necessary abilities to generate assessments of superior quality that precisely evaluate student learning. Efficient assessments not only measure comprehension but also guide teaching.

*Students.*Tests that are well-designed contribute to an impartial and precise assessment of student learning. This minimizes the likelihood of students facing repercussions due to errors in the text structure or ambiguous questions. The study's results have the potential to enhance the process of test create.

*Parents.*Well-designed tests provide parents with a more accurate understanding of their child's abilities and areas for improvement. This facilitates more knowledgeable discussions regarding academic advancement and focused assistance inside the household. The study's contribution to enhanced test building enables parents to gain a deeper comprehension of their child's educational progress.

*Education Stakeholders.*Educational policymakers and curriculum developers can benefit from this study as it provides valuable information. By emphasizing efficient methods for constructing tests, it can guide the development of standardized tests and classroom assessments that are in line with educational objectives. The synergy between research and policymaking ultimately enhances the entire education system.

Future Researchers. This study can serve as a significant basis for future researchers in the field of educational assessment. Through the analysis of the merits and drawbacks of test construction procedures, researchers can pinpoint areas where tests are being improperly utilized or misunderstood. This insight can inform the creation of enhanced testing techniques that better align with student learning and minimize the risk of bias. Furthermore, the study's focus on

professional development underscores the necessity for instructors to possess a thorough understanding of test construction. This can provide valuable insights for future studies on developing training programs that provide instructors with the necessary expertise to design and implement assessments in a highly successful manner.

Definition of Terms

For the purpose of the study, the following terms were defined conceptually and operationally to enrich the understanding of the research:

Use. McMillan, J. H. (1997) defined the use of test construction as a process of developing and administering tests acceptably to evaluate the learning results of students.

Misuse. Popham, W. J. (2001) defined misuse of testing as tests that are not linked with learning objectives, that place an excessive emphasis on rote memory rather than comprehension, or that are prejudiced against certain student groups are examples of this phenomenon. Assessments that are not used properly can give an inaccurate image of what students are learning and can make it more difficult to teach effectively.

Abuse. Hopfenbeck, T. N. (2019) defined abuse of assessments that do not adequately measure the talents that are intended to be measured using evaluation methods that restrict learning and place an emphasis on rote repetition using culturally biased examinations, which places certain student populations at a greater disadvantage than most other student populations.

Test Construction. Is the process of designing assessments to measure students' knowledge and skills, ensuring alignment with learning objectives and maintaining validity and reliability. According to Nitko (2004), effective test construction involves creating clear, unbiased items and proper scoring rubrics. It emphasizes the importance of avoiding common errors like ambiguity and bias to ensure fair assessments, which can improve both teaching and learning outcomes (McMillan, 2013; Popham, 2017).

Research Design

A descriptive research design was utilized in this study. This approach explored how intermediate-level teachers in a particular school in Pampanga potentially use, misuse, and abuse test construction in educational assessment. According to Rojo-Ramos et al. (2020), descriptive research seeks to offer a comprehensive portrayal of a phenomenon or population at a particular moment, concentrating on attributes and actions without changing factors or proving causal connections. In general, a descriptive design allows the researchers to accurately record the use, misuse, and abuse of test construction within this particular educational institution.

A descriptive research design was chosen for this study because it is well-suited for examining and documenting specific phenomena in their natural context. This approach enabled the researchers to investigate the test construction practices of intermediate-level teachers, highlighting potential patterns of use, misuse, and abuse without altering the conditions or seeking causal relationships. By focusing on a detailed and systematic exploration, this design ensured that the study provided an accurate snapshot of the current practices in test construction, which is crucial for identifying areas needing improvement and establishing a foundation for future interventions.

The study employed a checklist indicator and content analysis to systematically assess the test papers. The checklist served as a structured tool to evaluate specific elements related to test construction, such as clarity, alignment with learning objectives, and fairness. Content analysis was used to examine the qualitative aspects of the test papers, such as the phrasing of questions and the distribution of cognitive levels. Together, these methods facilitated a thorough examination of the test papers, ensuring that both quantitative and qualitative aspects of the teachers' practices were captured and analyzed comprehensively.

Research Respondents

The judicious selection of research respondents is paramount to the validity and reliability of a study. The choice of participants directly influences the quality and generalizability of the findings. By carefully considering factors such as demographics, experiences, and knowledge, researchers can ensure that their data accurately reflects the target population. A well-selected sample not only enhances the credibility of the research but also contributes to the development of meaningful insights and evidence-based conclusions.

The study included intermediate-level teachers and master teachers in the subjects of mathematics and science as respondents from one selected school in Pampanga. The study utilized 24 teacher-made test papers, each accompanied by its corresponding table of specifications (TOS), obtained from intermediate-level teachers. These test papers served as the primary source of data for the study. The 4th grade test papers contain 40 items each, while the 5th and 6th grade test papers contain 50 items each.

Sampling Technique

Considering the number of intermediate-level teachers in a single school was sufficiently limited to handle. This study utilized purposive sampling, which is a non-probability selection technique that selected documents needed. According to Obilor, E. I. (2023), "purposive sampling" is a technique that enables researchers to choose participants based on their assessment of who can provide the most

pertinent and illuminating information for the study issue. The population for this study consisted of teacher-made tests from teachers at an intermediate level in a specific school that had been deliberately selected in the Pampanga region. The goal was to include a diverse variety of experiences within the intermediate teaching population. This technique was appropriate since the research aimed to comprehend the particular difficulties and methods associated with test construction that intermediate teachers encountered, and it was expected that all relevant themes within this specific document were thoroughly explored.

A study titled "Purposive Sampling: Complex or Simple?" by Campbell et al. (2020) demonstrates the efficacy of purposive sampling in educational research. The study employed purposive sampling to select individuals who held broad expertise related to the research subject. Similarly, this study intended to collect comprehensive and precise data on the practices of test creation within a specific educational context by concentrating on a carefully selected school and its intermediate teachers.

Research Instrument

This study utilized an adopted indicator checklist. The indicator checklist aligns with both the study and its primary focus, test construction. This indicator allowed for a standardized evaluation of each teacher-made test paper while also allowing space for capturing any unexpected insights that may emerge during the analysis.

According to Nasir L. (2010), the checklist has a purpose beyond being a mere administrative procedure forced onto unaware employees. It functions as a communication tool, facilitating connections between members of the interdisciplinary team and their intricate task at hand. Utilizing a checklist at the onset of a complex procedure enables professionals to effectively inquire about pertinent matters and promptly recognize and resolve any difficulties. Moreover, the checklist helped the researchers to have a constant and concrete analysis of the teacher-made tests.

All data gathered from the documents were observed with absolute confidentiality. Lastly, the researchers successfully gathered the data by analyzing the teacher-made test papers. The final result of the indicator checklist would be subjected to the researchers' analysis and interpretation.

Ethical Considerations

This study strictly adhered to rigorous ethical criteria that protected the privacy and anonymity of the teachers, students, documents, and the school itself involved. Before any data collection, the said participants were provided with informed consent, which explicitly detailed the aim and methods of the study, as well as their right to withdraw at any point. Every survey guaranteed anonymity and all identifiable information on printed test papers was removed before analysis. Strict confidentiality was kept at all stages of the study process and in the conclusion following RA 10173, generally known as the Data Privacy Act.

Statistical Treatment of Data

To analyze the data, a content analysis approach was employed. This method, as supported by Krippendorff (2018), allowed for the systematic categorization and interpretation of textual data. The collected essays were meticulously examined to identify and categorize errors. Frequency counts were calculated to determine the prevalence of each error type.

With all the pertinent data collected, tallied, and tabulated through content analysis, the researchers utilized descriptive statistical techniques particularly frequency and rank as the statistical tool to determine the distribution of all respondents in each category. Moreover, these described the construction of the teacher-made test.

Frequency, Percentage and Rank of the Analyzed Teacher-Made Test

Table 1

Indicators of proper alignment in test construction

Subjects	Indicators	Frequency of Error
Mathematics	The test items aligned with the curriculum/learning competencies	0
	There are between three to four alternatives	0
	"Negatively stated items are used sparingly or the word that makes the stem negative is highlighted and/or capitalized"	0
	There are no cues that accidentally identify the correct answer	0
	There is no item that reveals that answer to the other item	0
	All distracters appear plausible	0
	"Alternatives are homogenous and nearly equal in length; if NOT of the same length they are arranged according to their length "	0
	"The "none of the above" or "all of the above" as alternatives are used sparingly"	0
	The reading level is appropriate to the test takers	0
	The use of "always" and "never" in the alternatives were limited	0
	Using the exact phrasing from the text is avoided	0
	"The test is organized in a logical manner; terms are arranged alphabetically; quantities are arranged in ascending or descending value/order"	0

International Journal of Academic Multidisciplinary Research (IJAMR)
ISSN: 2643-9670
Vol. 10 Issue 7 July - 2026, Pages: 491-503

	length; if NOT of the same length they are arranged according to their length "	
	"The "none of the above" or "all of the above" as alternatives are used sparingly"	0
	The reading level is appropriate to the test takers	0
	The use of "always" and "never" in the alternatives were limited	0
	Using the exact phrasing from the text is avoided	0
	"Standard font size was followed (Grades I-II font size is 14; Grades III-VI font size is 12)"	0
	"The required number of test items per grade level was followed (Grades I-II 30 items, Grades II-IV 40 items, Grades V-VI 50 items)"	0
	Table of Specifications (TOS 2-way) is congruent to the test questions	0
Total		0

The aforementioned table 1 for the subject math, illustrates the proper indicators of test construction in the realm of educational assessment, highlighting favorable aspects of its development. The use of test construction plays a vital role in ensuring that assessments effectively measure student learning and align with educational objectives. Properly constructed tests help educators accurately assess whether students have mastered the content and identify areas where further instruction may be needed. The data presented underscores the importance of aligning test items with curriculum competencies, ensuring that alternatives are clear, plausible, and free from unintentional clues. According to Nitko and Brookhart (2011), well-constructed test items should be aligned with curriculum standards, free of bias, and appropriate for the students' reading level. Moreover, the data reflects Haladyna's (2004) emphasis on using negatively stated items sparingly and ensuring that alternatives are homogenous in length to prevent confusion. These principles are essential in elementary education, where assessments guide both teaching and learning processes. The absence of errors across all indicators in the data further reinforces the importance of adhering to these best practices in test construction, ensuring that assessments are both fair and effective in evaluating students' knowledge.

The aforementioned table 1 for the subject science, illustrates the use of test construction in the realm of educational assessment, highlighting favorable aspects of its development. Effective test construction in science education is crucial for accurately assessing students' understanding of complex scientific concepts. The data indicates a strong alignment of test items with curriculum competencies, a critical factor that ensures relevance and validity in assessments. According to Gronlund (2006), aligning assessments with learning objectives is essential for accurately measuring student achievement. Additionally, the data highlights the importance of clear and fair alternatives, which are brief, homogenous in length, and free from unintentional clues—key practices supported by Haladyna (2004) for maintaining the validity of multiple-choice tests. The avoidance of negatively stated items and limiting the use of absolutes like "always" and "never" further enhances clarity, preventing student confusion, as noted by Nitko and Brookhart (2011).

The data also reflects consistency in the use of test construction, such as the adherence to standard font sizes and the number of items per grade level, ensuring uniformity, which Popham (2008) emphasizes as important for minimizing distractions during assessments. Finally, the congruence between the Table of Specifications (TOS) and test questions ensures a balanced assessment, covering all necessary content areas, which is fundamental for a comprehensive evaluation of students' scientific knowledge.

Table 2

Common errors found from math and science

Subjects	Indicators	Frequency of Error	Rank
	[illegible]	57	1
	[illegible]	11	2
	[illegible]	4	3
	All alternatives are grammatically correct relative to the stem	6	3
Science	The item stem clearly states the problem or question.	27	1
	The items are clear and easy to read	6	2
	All distracters appear plausible	2	3.5

In Table 2, some indicators are categorized as common errors of test construction in educational assessment. It may refer to more serious issues that undermine the validity and fairness of the test, requiring substantial changes to the test construction process to address. Moreover, analysis of the data reveals a cumulative total of 99 recurring errors detected throughout the first, second, third, and fourth quarters. The

occurrence of 99 common errors across an entire academic year indicates a significant issue with test construction. Moreover, these occurrences are under the subject of mathematics.

The data indicates that most of the issues involve errors such as items not being clear and easy to read, improper use of negatively stated items, and grammatical errors in alternatives. These issues can often be attributed to oversight or lack of training in proper test construction. Some errors that fell under a high frequency rate was the "The item stem clearly states the problem or question" which suggests a significant problem in how questions were framed. This is a critical component of test construction, as unclear stems can lead to confusion and misinterpretation, affecting the validity of the test. The test's logical organization, also being flagged, though less frequently, suggests occasional deliberate negligence or a lack of thorough review processes.

Research on educational assessment highlights the importance of clarity and proper construction in test items. Haladyna, Downing, and Rodriguez (2002) emphasize that well-constructed items should clearly present a problem and that all alternatives should be plausible to avoid misleading students. Furthermore, the use of clear and simple language is crucial to ensure that students understand the questions being asked (Downing, 2005).

In the analysis of test construction indicators for science, it was identified that several critical common errors were frequently made, potentially undermining the integrity and effectiveness of assessments. Notably, the common errors of grammatically wrong alternatives was observed, suggesting a prevalent issue in ensuring coherence and consistency between the stem and its associated options. This misalignment could lead to unintended cues or biases, potentially influencing students' responses and compromising the validity of the assessment.

Analysis of the data reveals a cumulative total of 40 recurring errors detected throughout the first, second, third, and fourth quarters. The occurrence of 40 common errors across an entire academic year indicates a significant issue with test construction. Moreover, these occurrences are under the subject of science. Some of the common errors are like distracters wherein it emerged as another concerning trend in test construction, indicating a failure to adequately challenge students' understanding and discernment.

These distracters, while intended to gauge comprehension, may inadvertently lead to confusion or uncertainty, thus detracting from the assessment's reliability. Additionally, the other type is logical organization, such as the absence of alphabetical or sequential ordering of terms and quantities, further contributing to the overall misuse observed in test construction. Such disorganization could impede students' navigation through the assessment, potentially affecting their performance and the validity of the results obtained. These findings underscore the critical importance of adhering to best practices in test construction to ensure assessments accurately measure student knowledge and skills.

The study titled "Test Construction Principles Violated in Selected Summative Tests in English" by Cristobal, J.- ar. (2022) also supports the idea that tests school teachers did not adhere to test construction principles. The study found that the majority of vocabulary items violated the same word class principle, while grammar items used a combination of different options. The majority of reading comprehension questions were answered based on prior knowledge, which may have impeded accurate and successful assessments.

Table 3

Data that may raise concerns ethically

Subjects	Indicators	Frequency of Error
Mathematics	The item stem clearly states the problem or question.	57
	The items are clear and easy to read	21
	There is only one alternative that is correct or represents the best answer	11
	The alternatives are brief and arranged for efficient scanning	4
	All alternatives are grammatically correct relative to the stem	6
Science	The item stem clearly states the problem or question.	27
	The items are clear and easy to read	6
	All distracters appear plausible	2
	"Negatively stated items are used sparingly or the word that makes the stem negative is highlighted and/or capitalized"	2
	All alternatives are grammatically correct relative to the stem	2
	"The test is organized in a logical manner; terms are arranged alphabetically; quantities are arranged in ascending or descending value/order"	1
Total		139

A predominant error identified in the analysis was the item stem's failure to clearly articulate the problem or question. This issue was notably prevalent and ethical concerns may arise, occurring in 57 instances within the dataset. Such ambiguity can significantly hinder students' comprehension of the task, potentially impacting their performance and the overall reliability of the assessment.

These errors affect the test construction as the improper use of negatively stated items, can lead to confusion and misinterpretation, as students might overlook negative terms, especially if they are not highlighted (Tarrant, Knierim, Hayes, & Ware, 2006). Ensuring that all alternatives are grammatically correct and plausible helps maintain the reliability of the test, as poorly constructed alternatives can

provide unintentional clues to the correct answer or confuse students (Haladyna & Rodriguez, 2013).

The high frequency of issues, such as unclear item stems, was particularly concerning where it affected clear communication of the problem in the stem in fundamental assessing the intended knowledge accurately. Misleading or confusing stems can invalidate the results and undermine the test's purpose (Haladyna, Downing, & Rodriguez, 2002). Proper logical organization of the test is also vital, as it affects the cognitive load on students and their ability to navigate the test effectively (Sweller, Ayres, & Kalyuga, 2011).

Among the common errors identified, one particular issue stands out: the item stem frequently fails to clearly state the problem or question. This recurring error was observed in 27 instances, significantly impacting the clarity and effectiveness of the test items. Moreover, this lack of clarity can lead to student confusion, potentially affecting their performance and the overall validity of the assessment.

These errors pose ethical concerns wherein the quality of test item construction that aligns with previous research highlighting the significance of clear, coherent test construction to maintain the validity and reliability of assessments. For instance, Haladyna (2004) emphasizes the importance of logical organization within tests, arguing that disorderly arrangements can hinder students' comprehension and compromise the validity of the assessment results. Similarly, Downing and Haladyna (2006) stress the necessity of plausible distracters in multiple-choice items, cautioning against the inclusion of overly obvious or implausible options that may lead to guessing rather than genuine understanding. Furthermore, Smith (2010) discusses the importance of grammatical consistency between item stems and alternatives, emphasizing the need for precise wording to minimize ambiguity and ensure fair assessment practices.

Intervention Program

An **intervention program** can serve as a targeted solution to address gaps in elementary teachers' test construction skills. By offering structured workshops, hands-on training, and continuous mentoring, the program aims to enhance teachers' understanding of test validity, reliability, and alignment with learning objectives. This initiative may also incorporate modules on ethical considerations, reducing errors, and practical applications of assessment principles.

Professional development can be done to effectively enhance elementary teachers skills such as creating a “Test Construction Skills Toolkit” wherein it will provide access to item banks and scoring rubrics. Incorporate technologies such as test generation software, online assessment platforms, and digital item response theory (IRT). Furthermore, these intervention programs can be used to create more efficient and reliable tests.

The data analysis from the previous tables reveals significant issues in test item construction that affects it, particularly in the clarity of item stems, which may categorized as abuse due to its high frequency and potential to undermine the assessment's validity. Interventions mentioned from the previous discussion to address errors are crucial for developing reliable, valid, and fair assessments. These findings are consistent with established research, which emphasizes the need for clear, well-constructed test items and logically organized tests.

Summary of Findings

The following findings, which addressed the objectives that were established, were obtained as a result of the study:

1. Test Construction in Elementary Education

- The data for the mathematics test construction show that all 15 key indicators have been properly applied, as indicated by the value of 0 for each item. These indicators include alignment with the curriculum and learning competencies, appropriate number of alternatives, minimal use of negatively stated items, and avoidance of cues that reveal correct answers. The data also show that no items inadvertently reveal answers to other questions, all distracters appear plausible, and alternatives are either homogenous or arranged logically. Additionally, there is limited use of "none of the above" or "all of the above" options, appropriate reading levels, and restraint in using extreme terms like "always" and "never." The test avoids exact phrasing from the text, follows standard font sizes, adheres to the required number of test items, and ensures that the Table of Specifications (TOS) aligns with the test questions Nitko and Brookhart (2011).

- The data for science test construction reveal a strong adherence to best practices, with all indicators properly used Haladyna (2004). Specifically, the test items are well-aligned with the curriculum and learning competencies (0 errors). There are between three to four alternatives for each question (0 errors), and these alternatives are brief and efficiently arranged (0 errors). The tests also ensure that only one alternative is correct (0 errors), with no cues accidentally identifying the correct answer (0 errors). Additionally, no items reveal the answers to other questions (0 errors), and the alternatives are homogenous and nearly equal in length (0 errors). Further, "none of the above" or "all of the above" options are used sparingly (0 errors), and the reading level is appropriate for the test takers (0 errors). The use of absolute terms like "always" and "never" in alternatives is limited (0 errors), and exact phrasing from the text is avoided (0 errors). The standard font size for different grade levels is followed (0 errors), and the required number of test items per grade level is met (0 errors). Lastly, the Table of Specifications (TOS) is congruent with the test questions (0 errors).

2. Common Errors in Test Construction by Elementary Teachers

The data for mathematics reveal that a majority of the items do not effectively present the problem or question, as indicated by 57 responses. Clarity and readability are not well-addressed, with 21 items being noted as not easy to read. However, fewer items strictly adhere to the requirement of having only one correct or best answer, with only 11 responses reflecting this criterion. The brevity and efficient arrangement of alternatives received minimal emphasis, with just 4 items meeting this standard. Similarly, only 6 items were reported to be grammatically incorrect in relation to the stem.

In contrast, the data for science show that while 27 items do not clearly state the problem or question, other aspects of the test design were less frequently met. Only 6 items were deemed not clear and easy to read, and even fewer (2 items) had no plausible distracters or not used negative statements sparingly with proper emphasis. The grammatical correctness of alternatives relative to the stem and the logical organization of the test were also minimal, with only 2 and 1 item(s) meeting these standards, respectively. These findings suggest that while the clarity of the question stem is prioritized in both subjects, other critical aspects of test construction, such as readability, plausibility of distracters, and grammatical correctness, need more attention, particularly in science (Sweller, Ayres, & Kalyuga, 2011).

3. Ethical Concerns in Test Construction Practices

The data for math and science reveal several ethical concerns related to potential misuse or abuse of test construction practices. In math, 57 items do not clearly state the problem or question, but only 21 items are not clear and not easy to read. Furthermore, only 11 items have a no or more single correct answer, raising concerns about the validity of the assessment. The low number of items (4) with lengthy and not efficiently arranged alternatives and the limited grammatical correctness (6 items) suggest that students may struggle to interpret the questions, potentially leading to unfair assessments.

In science, similar issues arise, with 27 items not clearly stating the problem or question but only 6 being not clear and not easy to read. Additionally, only 2 items feature no plausible distracters, and just 2 items not appropriately handle negatively stated stems. The organization of the test is also lacking, with only 1 item not arranged logically, which could confuse students and impact their performance Haladyna (2004).

4. Intervention Needs in Test Construction for Elementary Teachers

The data reveal critical areas where professional development is needed to improve elementary teachers' test construction skills. In math, there is a need for clearer and more readable test items, as only 21 out of 57 items meet this standard. Teachers also need training in creating questions with a single correct answer Nitko and Brookhart (2011), as only 11 items meet this criterion, and in writing concise and grammatically correct alternatives, with only 4 and 6 items respectively achieving these standards. In science, similar issues are present. While 27 items clearly state the problem or question, only 6 are clear and easy to read. Professional development should address the need for plausible distracters (only 2 items) and proper handling of negatively stated items (2 items). Additionally, only 1 item is organized logically, highlighting the need for training in structuring assessments effectively.

Conclusion

The analysis of the data suggests that there were substantial shortcomings in the construction of test items in both mathematics and science assessments. The primary issues identified include:

1. Test Construction in Elementary Education

The findings demonstrate that the test construction for mathematics among elementary teachers has been effectively executed, with all 15 indicators met according to the provided criteria. This suggests that teachers are proficient in creating assessments that align with educational standards, ensuring clarity, fairness, and appropriate difficulty levels for students. The proper use of test construction techniques reflects a strong understanding of assessment design, which is crucial in accurately measuring student learning and supporting their academic development. No major errors or oversights were detected, indicating that the test construction practices in this instance align well with best practices in education. While, the findings indicate that the test construction process for science is thorough and aligns well with educational standards and best practices. All aspects of the test, from alignment with curriculum competencies to formatting and readability, are effectively implemented. This reflects a well-constructed assessment that is likely to accurately measure student learning outcomes. The zero indicators of error suggest that the elementary teachers have a strong grasp of test construction principles and are applying them correctly in their assessments.

2. Common Errors in Test Construction by Elementary Teachers

Based on the data gathered, common errors in test construction among elementary teachers include a lack of clarity in the presentation of items, as seen in both mathematics and science. For instance, only a small number of items are easy to read (21 in mathematics, 6 in science),

indicating that clarity is often overlooked. Another frequent error is the failure to ensure that there is only one correct answer, with only 11 items in mathematics meeting this criterion. Additionally, teachers may struggle with organizing alternatives in a concise and efficient manner, as only 4 mathematics items adhere to this standard. In science, the limited use of plausible distracters and the minimal focus on logical organization of the test are also notable shortcomings.

3. Ethical Concerns in Test Construction Practices

The findings underscore the urgent need for targeted professional development to address key deficiencies in elementary teachers' test construction skills. The data reveal significant gaps in areas such as clarity, readability, grammatical correctness, and test organization. Teachers require training to create clear and accessible test items, ensure that each question has a single correct answer, and design plausible distracters. Additionally, improving skills in handling negative phrasing and logically organizing test items is crucial. By focusing on these areas, professional development can help teachers produce more effective, fair, and reliable assessments, thereby enhancing the overall quality of education and better supporting student achievement.

Recommendations

Based on the findings, the following recommendations are proposed to improve test construction practices:

1.Implement an integrated test design software like Formative that uses AI to evaluate the alignment of test items with educational standards, providing real-time feedback on clarity, difficulty, and fairness. This technology can help streamline the creation of assessments and ensure adherence to best practices without extensive manual review, thus maintaining high standards in test construction.

2.Utilize a tool such as ExamSoft that includes features for item analysis and readability scoring. This software can automatically flag issues such as unclear questions and inconsistencies in answer options, offering recommendations for improvement. This approach allows teachers to quickly address common errors and enhance the quality of their assessments with minimal additional effort.

3.Integrate Data and Analytics Tools such as Pearson's TestNav. This platform includes analytics features that track and evaluate test item performance and fairness across different student demographics. By using these analytics, educators can identify potential biases, assess the validity of distracters, and ensure that test items are equitably challenging for all students. The data-driven insights from TestNav can help in revising and improving test items to prevent unfair assessments and ensure accurate measurement of student abilities.

4.Leverage Khan Academy's Teacher Development Platform, which offers interactive modules on test construction and item writing. This platform provides targeted, technology-driven training that focuses on clarity, readability, and effective test design. Teachers can benefit from practical, hands-on learning experiences and immediate feedback, making professional development both efficient and impactful.

5.Future researchers could delve deeper into the specific challenges educators face when constructing assessments. By conducting qualitative studies, such as interviews and observations, researchers can gain valuable insights into the root causes of these difficulties. This could involve exploring factors like a lack of training, inadequate resources, or systemic pressures that hinder the development of high-quality assessments. Understanding these underlying issues can inform the development of targeted interventions and professional development programs to enhance educators' assessment creation skills.

By addressing these areas, educational institutions can enhance the validity and reliability of their assessments, ensuring that they more accurately measure student knowledge and skills.

REFERENCES

[1] Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. InternationalJournal of Evaluation and Research in Education, 9(1), 109-119.

[2] Ahyan, S., & Prabawanto, S. (2019, November). Students' Errors In solving the PISA mathematics problem using Newman's error categories. In Journal of Physics: Conference Series (Vol. 1363, No. 1, p. 012088). IOP Publishing.

[3] Barrientos, D. D. (2023). Assessing the knowledge of teachers in objective test construction procedure in the Teacher Education Programs. International Journal For Multidisciplinary Research, 5(5).

[4] Butakor, P. K. (2022). Using classical test and item response theories to evaluate psychometric quality of teacher-made tests in Ghana. European Scientific Journal, ESJ, 18(1), 139-168.

[5] Cahapay, M. B. (2020). Reshaping assessment practices in Philippine teacher education institutions during the coronavirus disease 2019 crisis. Pedagogical Research, 5(4).

[6] Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., ... & Walker, K. (2020). Purposive sampling: complex or simple? Research case examples. Journal of research in Nursing, 25(8), 652-661.

[7] Cristobal, J.- ar. (2022). Test Construction Principles Violated in Selected Summative Tests in English. American Journal of Arts and Human Science,

[8] Dela Rama, P. A. . (2022). The Conformity of Test Construction of the Achievement Test Papers of College Teachers: A Case Study. Silliman Journal, 52(2).

[9] Downing, S. M. (2005). Twelve steps for effective test development.

[10] Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. Applied Measurement in Education, 15(3), 309-334.

[11] Haladyna, T. M., & Rodriguez, M. C. (2013). Developing and validating test items. Routledge.

[12] Inko-Tariah, D. C., & Okon, E. J. (2019). Knowledge of test construction procedures among lecturers in Ignatius Ajuru University of Education, Port Harcourt, Nigeria. Academic Research International, 10(1), 130-138.

[13] Izard J (2005). Overview of test construction. In K. N. Ross. Quantitative Research Methods in Educational Planning. Paris: UNESCO International Institute for Educational Planning.

[14] Krippendorff, K. (2018). Content analysis: An introduction to its methodology.Sage Publications.

[15] Kumar, D., Jaipurkar, R., Shekhar, A., Sikri, G., & Srinivas, V. (2021). Item analysis of multiple choice questions: A quality assurance test for an assessment tool. Medical Journal Armed Forces India, 77, S85-S89.

[16] McMillan, J. H. (1997). Classroom Assessment. Principles and Practices For Effective Instruction. Allyn & Bacon, A Viacom Company, 160 Gould St., Needham Heights, MA 02194; Md Yusof, Sallehuddin. (2023). The Six Stages of Test Construction.

[17] Monica, S. (2021, March). The challenges in constructing and administering an objective test for beginners' students during the COVID-19 pandemic. In International Conference on Educational Sciences and Teacher Profession (ICETeP 2020) (pp. 441-446). Atlantis Press.

[18] Nasir L. (2010). The Checklist Manifesto: How to Get Things Right. London Journal of Primary Care, 3(2), 124.

[19] Neuendorf, K. A. (2016). The content analysis guidebook. Sage Publications.

[20] Obilor, E. I. (2023). Convenience and purposive sampling techniques: Are they the same. InternationalJournal of Innovative Social & Science Education Research, 11(1), 1-7.

[21] Opateye, Johnson. (2023). Assessment of Teachers' Knowledge of Essay and Multiple-choice Tests Construction in Secondary School Chemistry in Bwari Area Council, Abuja. 10. 72-81.

[22] Opie, O. N., Oko-Ngaji, V. A., Eduwem, J., & Nsor, A. J. (2021). An assessment of science teachers' utilization of the knowledge of test construction procedure in multiple choice objective tests in secondary schools in Yala LGA,Cross River State. British Journal of Education, 9(11), 54-62.

[23] Popham, W. J. (2001). The truth about testing: An educator's call to action.ASCD.

[24] Quansah, F., Amoako, I., & Ankomah, F. (2019). Teachers' test construction skills in Senior High Schools in Ghana: Document analysis. International Journal of Assessment Tools in Education, 6(1), 1-8.

[25] Rojo-Ramos, J., Ferrera-Granados, C., Fernández-Guerrero, M., Manzano-Redondo, F., Garcia-Gordillo, M. Á., Polero, P., & Adsuar, J. C. (2020). A descriptive study on the training and attitude of future teachers towards educational inclusion. Sustainability, 12(19), 8028.

[26] Rutkowski, L., & Rutkowski, D. "A call for a more measured approach to reporting and interpreting PISA results." Educational Researcher, 45.4 (2016): 252–257.

[27] Sjøberg, S. (2020). The PISA-syndrome-How the OECD has hijacked the way we perceive pupils, schools and education. Confero: Essays on Education, Philosophy and Politics, 7(1), 34-88.

[28] Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory. Springer.

[29] Tarrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. Nurse Education in Practice, 6(6), 354-363.

[30] Therese N. Hopfenbeck (2019) The use and abuse of assessment, Assessment in Education: Principles, Policy & Practice, 26:6, 637-642,

[31] Zijlmans, E. A. O., Tijmstra, J., van der Ark, L. A., & Sijtsma, K. (2019). Item-Score Reliability as a Selection Tool in Test Construction. Frontiers in Psychology, Volume 9.

1. Respondents Demographic Profile

2. Level of the respondents' attitude, perceived skills, knowledge, and Practices towards educational assessment.

The educational assessment attributes of the higher education institution's faculty

al of A Research (IJAMR)
026, P

Ass the Higher Education Institutio

Reymond Q. Fajardo[1], Mark Aron S. Mactal [2]

[1]Santo Tomas Campus, Pampanga State University, Pampar
rqfajardo@pampangastateu.edu.ph

[2]Santo Tomas Campus, Pampanga State University, Pampanga,
masmactal@pampangastateu.edu.ph

Abstract: *The high demand for quality education today comes with the high demand for clear-cut educational assessment methods. The illiteracy in educational assessment is a suicide to the system. Hence, it is essential for educators, specifically those teaching in tertiary level, to have enough knowledge in this aspect. This study explored the practices of HEI's faculty towards educational assessment with regard to attitude, perceived skills, knowledge and practices. Using a standardized questionnaire by Alkharusi, et. al., (2009), the results from 103 educators revealed that teachers have neutral attitude towards the educational assessment and that they perceived themselves skilled in it. However, only one (1) respondent is emerged competent. In terms of practicing the use of reliable means in student assessment, respondents were rated "moderate". Analysis of variance (ANOVA) revealed that there are no significant differences among the profile of the respondents with their corresponding attitude, perceived skills, knowledge, and practices towards educational assessment. Based on the findings, it is highly recommended that the university offer trainings and seminars for the competence of faculty members in educational assessment. As fundamental part of professional growth, instructors, regardless of specialization must be engaged in teacher development programs concentrating in educational assessment.*

Keywords—*Educational Assessment, Attitude, Self-Perceived Skills, Assessment Knowledge, Assessment Practices*

1. Introduction

Ignorance about educational assessment is a risk in students' learning. Fuller (2019) defined educational assessment as a systematic collection, review and application of information regarding educational programs undertaken. It uses tools and instruments for the educator to measure, evaluate, and document and report the academic readiness, the learning progress, the acquisition of skills, and/or the educational needs of students to improve the teaching-learning process. As Popham (2009) argued, an educational assessment is a basic yet a vital and a critical component in the teaching-learning process. However, appropriate assessment of the key competencies and skills of students remains a major challenge that educators face in today's teaching Siarova et.al. (2017). Despite the significant role of tests and appropriateness of assessment strategies and techniques, teachers still struggle in classroom assessment. The evaluation goes beyond student achievement since it must consider all aspects of the teaching-learning process. This yields to looking for best educational decisions based on the results of other forms of assessment (Jabbarifar, 2009). Furthermore, in higher education instruction, the process of assessment is more influential to students' learning than the lecturer's teaching (Clarence et. al., 2017).

In Philippines setting, assessment is one of the skills that a teacher must have to be considered competent, and it is grounded on the Philippine Professional Standards for Teachers or PPST 2017. Nevertheless, a deeper understanding of the suitability of teachers' competencies and macro-skills in student assessment which guides teachers to be better practitioners of the profession is a prerequisite (Magno 2013). Although, the university mentors are given much autonomy in teaching, there are still gaps with regard to the skills in classroom assessment practices. As Barandon and Atian (2016) claimed in their study, the faculty members have limited awareness in the assessment techniques which results to poor utilization of the training. Moreover, this was supported by the study of Gonzales (2014), which said that most of the college instructors are not fully oriented to the significance of assessment in the pr ic, Lasate led in hi oach ave hard e preparing for the test including testing and te lated p rs. A onally, at teach adic a ne or training assessment in general.

This study fo d on the s orientation assessment specificall suring attitude, knowledge, perceived skills, and pr es of facu ver, it also p de rd's eye view of the institution provide quality assurance, whether it is to uphol hig quality in actic luding th ility ovi ality assessment methods to the students.

The primary purpose of the study was to evaluate the educational assessment attributes of the higher education institutions' faculty. The research results shall serve as a reference point or basis in identifying the need for capability building for professional growth of faculty members. Darling-Hammond, et.al (2017) defined professional development in education as basis