

# Teacher's Test Construction Competencies on Student's Academic Performance in Science

A Graduate Thesis  
Presented to the Faculty of the Institute of Education  
Bulacan Agricultural State College  
San Ildefonso, Bulacan

In Partial Fulfillment  
of the Requirements for the Degree Master of Arts in Education  
(Major in General Science)

**<sup>1</sup>Princess Milarpiz Pastrana**  
December 2022

Pastrana, Princess Milarpiz.: Teachers' Test Construction  
Competencies on Students' Academic Performance In Science  
Adviser: Dr. Lauro S. Nuñez

**ABSTRACT**

**This study determined the influence of teachers' test construction competencies on the students' academic performance in Science during the School Year 2022-2023. With explanatory sequential as research design and Science teachers, students, and school heads as respondents of the study. Findings showed that majority of the teachers belong to the age group of 24 to 38 years old, master's degree holders, female, Biology major, and attended seminars in test construction. The Science teachers obtained a verbal description of "almost always" to "always" and believed that they are competent in test construction, as validated by their school heads. The academic performance of the junior high school students in Science was described as "satisfactory". Based on the findings of the study, the following conclusions were drawn: Science teachers had a greater assessments of themselves as compared to their school heads in test construction competencies in terms of developing assessment methods appropriate for instructional decisions. Science teachers who are female, finished master's and doctorate degrees, and attended seminars in test constructions had greater level of test construction competencies in terms of administering, scoring, and interpreting results. Teachers' test construction competencies are positively correlated to students' academic performance in Science.**

## CHAPTER ONE

### THE PROBLEM AND ITS BACKGROUND

#### A. Introduction

Among other areas of teaching, student assessment is an indicator of a teacher's effectiveness in carrying out learning goals. Assessment is the process of gathering information to aid teachers in making educational decisions about students through determining and evaluating their needs and lacks. It allows students not only to know the learning goals but also to acknowledge the need to achieve them. Also, this includes the diagnosis as to which level the students are placed in terms of their learning to determine the most effective ways or steps to attain learning goals. It is used as a basis to review the effectiveness not only of instructional materials but of the curriculum being implemented which consequently inform policies.

A "Standards for Teacher Competence in Educational Assessment of Students" is crafted to determine and address issues concerning student assessments. The standards include teachers being skilled in choosing and developing assessment methods appropriate for instructional decisions; in administering, scoring and interpreting the results of various assessment methods; in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement; in developing valid student grading procedures; in communicating assessment results to students, parents, co-educators, and other stakeholders; and in recognizing unethical, illegal, and inappropriate assessment methods and uses of assessment information (Shallianah, 2016). The alternative of the traditional assessment which is known as the assessment *of* learning puts emphasis on summative assessment which is administered at the end of a unit, a course, or a program through examination or tests, and feedback are provided in the form of marks or grades (Rahman & Majumber, 2015). On the other hand, assessment *for* learning puts emphasis on formative assessment where data from students' outputs in worksheets, responses during questioning or recitations, and observation as facilitated by teachers are collected. Assessments are planned, implemented, and evaluated through their purposes which are determined through looking into the differences between the two types of assessment: assessment *for* learning and assessment *of* learning.

Tests are found out to be the primary methods of assessment. There are three types of tests namely standardized tests, diagnostic tests, and teacher-made tests. Some of the reasons why most teachers use tests are assessment tools treated students equally and students are familiar with the method (Halinen et.al, 2014 as cited in Rawluysk, 2018). These are some of the reasons why some criticize testing as regards its capacity to reflect learning. Additionally, when principles of test constructions are violated, it could hamper effective and successful assessments (Cristobal, 2022). However, some educators argue how it positively impacts learning, most especially when teachers have the necessary competencies to formulate and administer these types of assessment tools.

The areas of competencies in effective student assessment teachers must possess include "determining the purpose of each testing exercise; giving specific, measurable objectives; creating a content outline and a test plan to serve as a guide for item construction; choosing appropriate test item formats; formulating items which are clear, precise, unambiguous, and with appropriate difficulty and discriminative indices; developing a marking guide which complement the test; performing item analysis; providing clear directions for test administration and test-taking; and conducting a test review to ensure no errors were made during item construction" (Chidolue, 1999, as cited in Amani et.al, 2021). These areas were used in developing the Teachers' Competence Questionnaire (TECOM-Q) in order to measure test construction competencies among teachers.

To strengthen classroom assessment for the K to 12 Basic Education Program and to adapt to the needs of the 21<sup>st</sup>-century learners, the Department of Education emphasizes the assessment of learning standards, namely: content standards, performance standards, and learning competencies that are outlined in the curriculum (DepEd Order No. 8, s.2015). Content standards identify what the students should know and understand. Performance standards are the abilities and skills that students are expected to show and demonstrate in relation to the integration of the content standards and 21<sup>st</sup> century skills of the students. Learning competencies refer to the knowledge, understanding, skills and attitudes that students need to demonstrate at the end of each lesson.

The Department of Education supports and promotes the professional development of teachers in all areas but especially in assessment. Taking into consideration the continuing professional development strategy for teaching and learning, the department started the implementation of "Learning Action Cell." DepEd Order NO. 35, s. 2016, mandates all teachers to form groups that will aid and collaborate to solve and enhance skills in all facets of teaching-learning process that will benefit the students. Each teacher should know and understand the importance of implementing learner-centered policies on assessment. Students should be informed on how their assessment results will be used and how they will be graded. Furthermore, bringing the 21<sup>st</sup> century skills into consideration and enhancing teaching and assessments with simple ICT integration is highly promoted.

In addition, the Department of Education released "Interim Policy Guidelines for Assessment and Grading in Light of the Basic Education Learning Continuity Plan" during the peak of the COVID-19 pandemic in 2020 to offer instructions on the evaluation of students' learning. DepEd reiterates the use of creative and innovative assessment tools. (DepEd Order No. 31, s.2020). Teachers should always ensure that assessment activities: align with the most essential learning competencies; are

reliable, valid and transparent; are fair, inclusive and equitable; are practical and manageable for both teachers and learners; give learners a range of ways to demonstrate their learning; and provide timely and accurate feedback. And although paper and pencil tests may seem a bit old and traditional, it is still proven to be effective (Hardcastle, et.al, 2017).

Teachers must be able to have a mastery of test construction in terms of language use, content coverage, item organization, and test guidance.

Language appropriateness is essential in test construction because the language being used, particularly its lexicon and syntax, determines how the students will comprehend the test directions and test items, which follows how they analyze and consequently, provide their answers. When the vocabulary used in test is too complex for the students, negative test emotions like frustration, fear, and anxiety may occur resulting in poor test performance. Therefore, language used in the tests should be consistent with the language being used in the classroom to make students confident in taking them. Also, test items should be simple, clear, without biases, non-functional material, and extraneous clues (Haladyna, 2018).

The content coverage allows students to demonstrate the competencies they have acquired in the lessons. The test must cover subject areas which were pursued during class discussions prior to the administration of the test. Also, the test must measure complex behavior like the ability to make practical applications instead of only measuring recall. Concepts that were given emphasis should have been reflected in the test. In retrospect, students tend to focus their attention on subject matter being highlighted the most by their teachers as they were implied to be the most important or the most challenging ones, and therefore will be reflected in the test. Relatively, students may find these subject matters as irrelevant if they found out that they were not included in the test as there is a mismatch of what was taught and what is being assessed (Ughamadu & Ifeyinwa, 2021).

Item organization concerns about how the test items presented in the same format are grouped together. It also deals with how the teachers provide sufficient time for students to complete the test.

Test guidance is observed before administering the test to where the teacher asks for feedback as regards its effectiveness through consulting with colleagues, editors, and testing specialists to check and provide their inputs on the appropriate difficulty of test questions, and the test format which may confuse or distract students.

Lack or inability to apply these test construction competencies will result to inaccurate assessment, and therefore impede students in reaching their full academic competence. Poorly designed tests may cause misleading feedback not only to the students, but also to the teachers as designing a learning reinforcement or intervention, and planning of the next lesson is heavily influenced by how the students performed in the preceding tests.

Considering that assessment of learning is the first among the five key components of a personalized learning model developed by Courcier (2007) as cited in Miracle (2021), test construction competencies should be developed, mastered, and efficiently employed by teachers as it will determine how the following components will be managed. The following components are effective teaching and learning strategies, curriculum entitlement and choice, school organization, and strong partnerships beyond the school.

Looking into the teachers' competencies establishes and develops testing practices leading to positive effects on the reliability and validity of test scores, which in turn, increases students' levels of academic performance.

Academic performance is the outcome of students' efforts in examinations. It is measured by grades from each quarter. It shows how well the students accomplish the tasks required by the lessons.

During the first few years of the adoption of the K to 12 Basic Education Program, assessing the content of the students' achievement in Science suffered due to poorly designed assessment tools and lack of sufficient resources (Berame, et.al, 2018). Test questions that were poorly crafted resulted in low students' academic performance, that is, even though the questions were created on their level of standard. The results of the study indicated that the students lack the prior knowledge to answer the questions. Students did not do well in their tests even though they have enough knowledge or are well-prepared. It was recommended that teachers use reliable and valid test questionnaires to become more suitable for 21<sup>st</sup> century science high school students.

Thus, this research examines the existing assessment techniques and practices which emerge from the test construction competencies among public high school teachers and how these competencies influence the academic performance in Science of high school students in the new normal.

#### *B. Statement of the Problem*

This study determined the influence of Grade 7 and 8 public high school teachers' test construction competencies on the students' academic performance in Science in public high schools in San Miguel, Bulacan during the School Year 2022-2023.

➤ *Specifically, it Sought Answers to the Following Questions:*

- *How may the socio-demographic profile of the Science teachers be described in terms of:*
  - ✓ Age;
  - ✓ Sex;
  - ✓ Educational Attainment;
  - ✓ Specialization;
  - ✓ *Seminars Attended Related to test Construction?*
- *How may the Test Construction Competencies of the Public High School Science Teachers as Assessed by the Teachers Themselves and by their Respective Department Heads be Described in Terms of:*
  - ✓ Developing Assessment Method Appropriate For Instructional Decisions
  - ✓ Administering, Scoring And Interpreting Results
  - ✓ Using Assessment Results
  - ✓ Developing Valid Students Grading Procedures
  - ✓ Communicating Assessment Results
  - ✓ Recognizing Unethical And Illegal Assessment Methods?
- *How may the high school students' academic performance in Science be described in terms of their average grade in the subject in the first grading period?*
- *Is there a significant difference between the assessments of the department heads and the Science teachers themselves as regards their test constructions competencies?*
- *Is there a significant difference between and among the Science teachers' test construction competencies when they are grouped according to their demographic profile?*
- *Is there a significant relationship between the teachers' test construction competencies and their students' academic performance in Science?*
- *How important are the teachers' test construction competencies on students' academic performance in Science?*
- *What program of activities can be crafted from the findings of the study?*

#### C. *Hypotheses*

The Following Hypotheses were Tested in the Study:

- ✓ There is no significant difference between the assessments of the department heads and the Science teachers themselves as regards their test constructions competencies.
- ✓ There is no significant difference between and among the Science teachers' test construction competencies when they are grouped according to their demographic profile.
- ✓ There is no significant relationship between the teachers' test construction competencies and their students' academic performance in Science.

#### D. *Conceptual Framework*

This study is anchored to classical measurement theory (CMT), which also provides an adequate model for most classroom tests. The classical measurement theory (CMT), which serves as an effective model for majority of in-class exams, is the bedrock of this study. This test development process model, according to Veldkamp (2015), is applicable to almost all types of cognitive tests, including those meant to measure achievement or ability, tests used for academic or employment placement, tests meant to identify learners' strengths and limitations, tests meant to give learners formative feedback on progress as well as a final summative measure of cognitive achievement in a domain, and more.

Assessment is a highly valued area of specialization teachers must account for. It mainly provides both teachers and students an effective feedback mechanism towards academic progress. The quality of the feedback and the efficiency of learners' acceptance, respond, and utilization are needed to be assessed to ensure that the learning objectives are attained. However, to promote effective assessment, teachers must have the necessary competencies required to deliver quality student assessments. Teachers' intellectual character of teaching is pronounced as a determining factor of teaching quality (Lowyck, 1994 as cited in Ibad, 2018). Their intellectual character is enriched with their ability to reason, deduce, make decisions, solve problems, and analyze. Therefore, it is important that becoming teachers master the procedures of test construction as were introduced in their Teacher Education course namely, Assessment for Learning for them to be armed with the essential knowledge and skills in student assessments.

A survey of conducted studies and literature seeks to deliver an assessment of a particular teacher’s competency. However, there is no unified comprehensive concept of teaching competence which may cover the diverse areas of competent performance. From this dilemma emerges the construct of the interpretive model of competent performance. The model begins with the base which includes the teacher’s knowledge, skills, attitudes, and personal characteristics, followed by his process of decision-making, actions, and consequences which is reflected on the product of the learning process.

Teachers should be knowledgeable as regards the strengths and weaknesses of multifarious assessment methods to pick which one is the most appropriate, adequate, and fair for a particular learning situation. Also, considering that assessment demands of the classroom change due to dynamic trends in the teaching-learning process, teachers must be able to make informed decisions in creating and implementing assessment methods which complement with the students’ competencies.

Test construction procedures include considering target learners’ age, ability, and gender, defining specific objectives, and learning outcomes, creating content outline, preparing a Table of Specifications, reviewing related textbooks, constructing clear, precise, and unambiguous test items, measuring how much time is sufficient in test-taking, creating clear test instructions, reviewing test to correct errors made during test construction, using clear language and appropriate format, preparing a guide, and performing item analysis (Amani et.al, 2021).

Furthermore, teachers’ decision-making is not only crucial at a classroom level as educational decisions define the actions they will be planning and implementing thereafter. This impacts the community to where the school is located, its school district, and thus, society, in its entirety.

The interpretive model of competent performance developed by Roelofs and Cito (2022) presents a framework on how to assess teachers’ professional competence. To accurately understand the model, the developers differentiate competence with competency. Competence is a comprehensive concept for abilities of a person which in this research focuses on teachers’ while competency is only a part of a person’s competence – it is used to label a particular ability. Person’s competencies are their integrated action proficiencies composed of knowledge structures; cognitive, interactive, emotional, and psychomotor skills; attitudes and values necessary for performing tasks, solving problems; and more generally, his abilities to function in a particular occupation, organization, position, or role.

Anchored to this model are the standards for teacher competence in educational assessment of students which to a great extent revolve around the importance of teachers’ decision-making. The following standards involve teachers being skilled in choosing and developing assessment methods appropriate for instructional decisions and being skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement (Shallianah, 2016).

Lastly, the consequence of the actions made by the teachers is examined on how they positively or negatively affected students’ learning. It is important for learners to understand the criteria and targets of the assessment process in order for them to direct their learning and development towards meeting and achieving them. In doing so, students become more invested and actively engaged in the process of recognizing how their academic performance relates to the criteria and standards (Brown, 2019).

Based on the presented framework and survey of relevant research and academic standards, the researcher developed the following paradigm that served as guide in the conduct of the study.

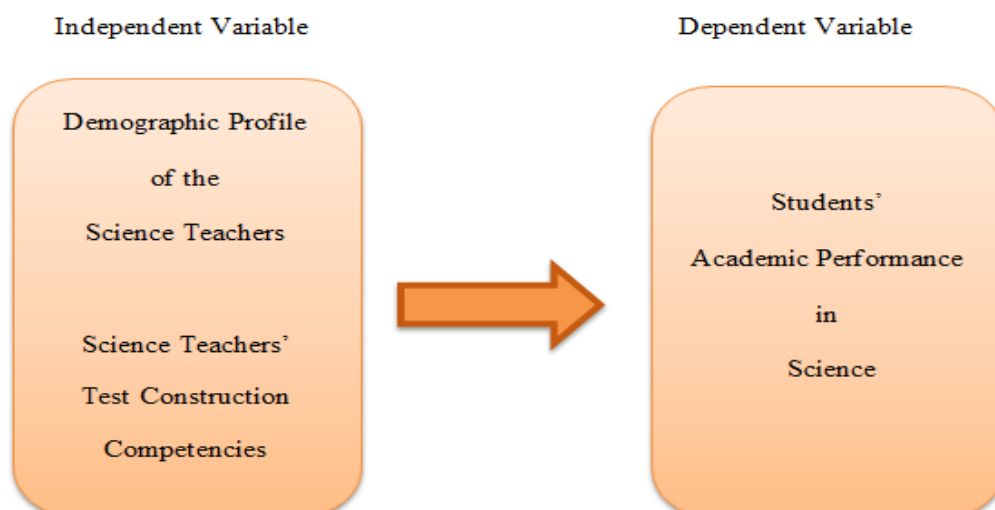


Fig 1 Paradigm of the Study

Figure 1 illustrates the paradigm of the study shows the relationship of the test construction competencies of public high school teachers and students' academic performance in Science. Moreover, in the figure, the independent variables are the demographic profile of the Science teachers. Additionally, the test construction competencies of public high school Science teachers is also considered as independent variable of the study. On the other hand, the dependent variable is students' academic performance in Science. It was hypothesized (as implied by the arrowhead) that a significant relationship existed between the aforementioned variables.

#### *E. Significance of the Study*

The study presents its significance in setting out defined procedures on how to conduct evaluation on teachers' competence in student assessment and look into solutions on how to make their competencies more adaptive to the learning demands and needs of the 21<sup>st</sup> century students.

**Public School Administrators.** The study will help them analyze the realities of student assessment as employed by the teachers. From the policymaking, these realities will aid administrators and educational leaders in reviewing, refining, and retooling the policies and standards of the implemented curriculum. This will guide them in gauging the test construction competencies among public high school teachers, training, and development. Programs will be packaged and carried out to them to address competence gaps.

**Teacher.** The high school teachers will be aware of the effectiveness of their test construction competencies, particularly as assessed with students' academic performance in Science. Also, it will direct them to what they need to learn and update as regards their test construction knowledge and modify their test construction skills considering the fast-changing trends in the academe.

**Students.** The results of study will increase students' level of academic performance in Science. As teachers' assessment tools will be modified this will become more complementary or responsive to the competencies acquired by the students. Learning will be measured with more efficiency and accuracy; thus, academic performance will produce the desired learning goals and standards.

**Future Researchers.** The study will serve as a reference for future researchers in their conduct of an in-depth survey of teachers' competencies, specifically on student assessment, therefore, enriching the pool of data highly essential for a much-improved educational system.

#### *F. Scope and Limitation of the Study*

In the conduct of this study, variables were limited to teachers' demographic profile, their test construction competencies, and their students' academic performance in Science. The demographic profile of the Science teachers was described only in terms of their age, sex, educational attainment, area of specialization and seminars attended related to test construction. Meanwhile, the test constructions competencies of the Science teachers focused only to developing assessment method appropriate for instructional decisions, administering, scoring and interpreting results, using assessment results, developing valid students grading procedures, communicating assessment results, and recognizing unethical and illegal assessment methods.

Additionally, the Science teachers' department heads or principal, in cases where the school does not have a Science department head, rated the teachers' test construction competencies using teacher-made tests as basis. The academic performance of public high school students was described in terms of their average grade in Science in the first grading period.

The respondents of this study were the Grades 7 and 8 Science teachers, Science department head teacher or school principal and students in all public high schools in San Miguel, Bulacan. The study was conducted in the first quarter of School Year 2022-2023.

#### *G. Location of the Study*

The study was conducted in all public high schools in San Miguel Bulacan. The schools involved in the study are San Miguel National High School, Vedasto R. Santiago High School, John J. Russell Memorial High School, Partida National High School, and Emilia Perez Ligon High School (formerly Maligaya High School).

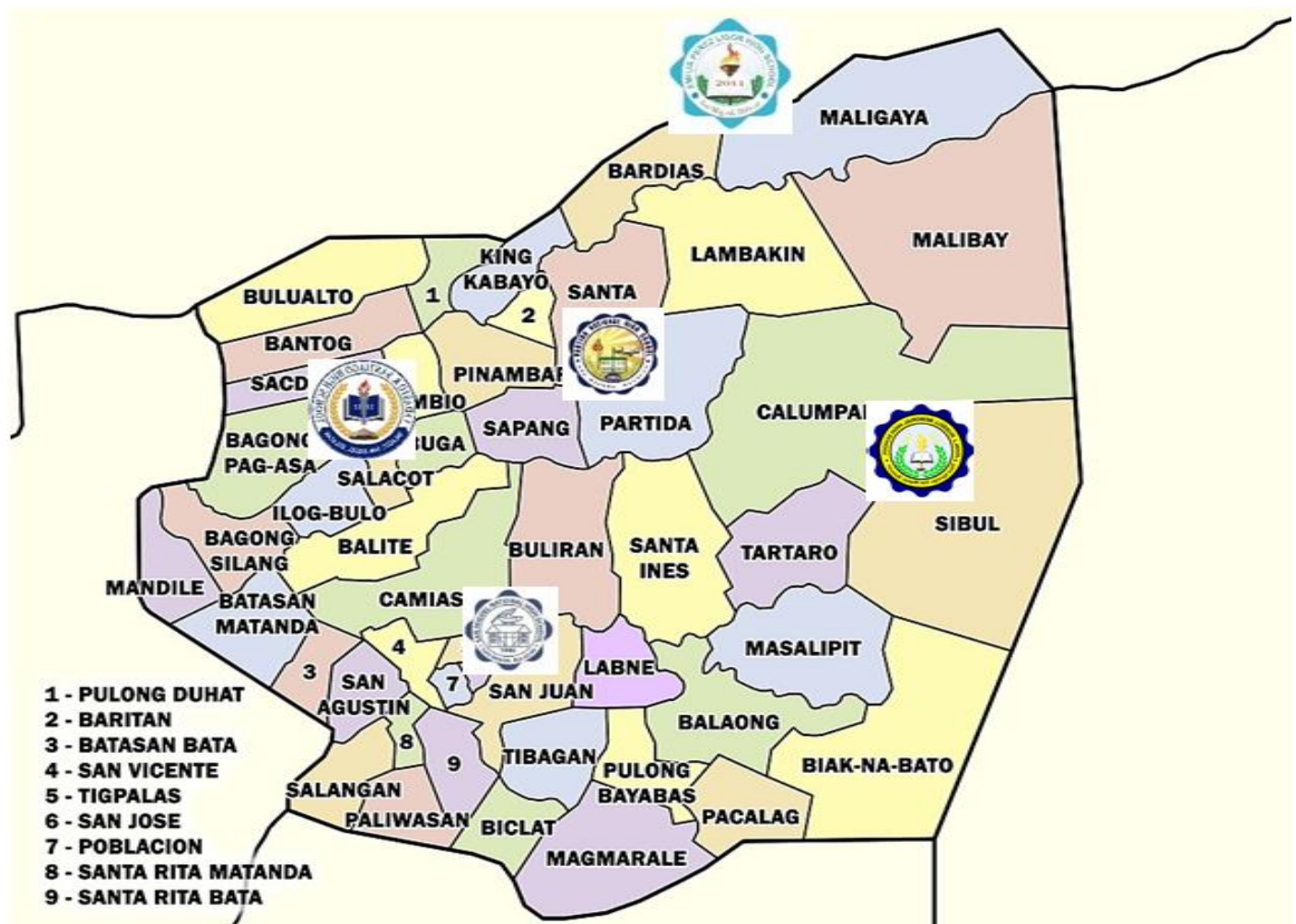


Fig 2 Map of San Miguel, Bulacan

Source: [https://en.wikipedia.org/wiki/San\\_Miguel,\\_Bulacan](https://en.wikipedia.org/wiki/San_Miguel,_Bulacan)

### ➤ Definition of Terms

- The following definitions are presented to explain how the terms are operationally used in the course of the study.
- Academic Performance. This refers to the evaluation of a learner's performance throughout the grading period as influenced by the teacher's test construction competencies.
- Administering, Scoring and Interpreting Results. This refers to teachers' test construction competency in relation to conducting exams, as well as the analysis and interpretation of examination data.
- Assessment. This refers to the systematic process used to evaluate, measure, and document students' academic performance and development.
- Assessment of learning. This refers to summative assessment where teachers make use of grades or marks as feedback on students' academic performance. It is used by teachers for the placement and promotion of their students.
- Assessment for Learning. This refers to formative assessment with the purpose of providing teachers information as regards students' academic progress which will enable them to determine what needs to be revised in planning instruction.
- Communicating Assessment Results. This refers to teachers' test construction competency which focuses on explaining the importance of taking examinations and conveying assessment results to students.
- Competence. This refers to the teacher's combination of knowledge, skills, values, and personal characteristics, enabling the teacher to act professionally and appropriately as regards test construction.
- Competency. This refers to a teacher's actual performance in test construction.
- Content Coverage. This refers to test construction competency which focuses on the relevance of the test items with the concepts or subject areas taught during class discussions.
- Developing Assessment Method Appropriate for Instructional Decisions. This refers to teachers' test construction competency as regards using assessment data effectively when making decisions about individual students, organizing instructions, and developing curricula.



- **Developing Valid Students Grading Procedures.** This refers to teachers' test construction competency which is mainly concerned in providing students with feedback so they can assume responsibility for their education and to provide information to those who support these learners.
- **Illegal Assessment Methods.** This refers to teachers' exam malpractices that are immoral and unacademical. This includes practices that teachers should avoid including but not limited to test paper or key to corrections leakage, falsification of exam results, selling of test papers, and tampering with exam results.
- **Item organization.** This refers to a test construction competency which shows how test items are placed in a systematized format.
- **Language Use.** This refers to providing test directions and questions that are clear, precise, and unambiguous.
- **Test.** This refers to an assessment tool developed by teachers or testing specialists that is given to students in order to measure a particular knowledge or skill as targeted by a learning objective.
- **Test Construction Competencies.** This refers to the knowledge and skills teachers have as regards formulating and administering reliable and valid tests.
- **Test Construction Skills Inventory.** This refers to a survey instrumentation used to measure teachers' competence in test construction.
- **Test Guidance.** This refers to teachers' test construction competency which is concerned on how tests are constructed in reference to textbooks or lectures used in class or through following a test blueprint.
- **Recognizing Unethical and Illegal Assessment Methods.** This refers to teachers' test construction competency as regards recognizing procedures that lead to misuse of assessment results and with the procedures that lead to overuse of assessment results.
- **Using Assessment Results.** This refers to teachers' test construction competency with respect to utilizing assessment results and accumulated assessment information to facilitate students' learning.

## CHAPTER TWO METHODOLOGY

The information about the research and sampling procedures utilized by the researcher are provided in this chapter. The research design that was employed, as well as the data gathering techniques, and data analysis scheme are also discussed.

### ➤ *Research Design*

The purpose of this research study is to examine the role of teachers' test construction competencies on the academic performance in Science among high school students. The research design follows the sequential explanatory design which incorporates quantitative and qualitative approaches in two consecutive phases within one study. This design capitalizes on the results of the first phase of the research which is the quantitative approach using survey questionnaire to draw insights for the second phase which is the qualitative approach using thematic analysis. The findings from both phases of the study were then examined and combined to draw the conclusions.

In the quantitative data collection phase of the study, the researcher collected survey data from the Science teachers and department heads in selected public high schools in San Miguel, Bulacan to assess their test construction competencies. Students' grades in Science for the first grading period were collected upon securing a signed consent.

During the qualitative data collection phase, through conducting a semi-structured interview, the researcher explored the teachers' and students' insights on the test construction competencies of Science teachers and how they affect students' academic performance in Science.

Upon gathering data both from the quantitative and qualitative phases, the last phase was facilitated by the researcher through analyzing convergence among the findings to form meta-inferences or an integrated understanding of the findings.

### ➤ *Data Gathering Techniques*

Before the conduct of the study, the researcher sought permission from the Schools Division Superintendent of Bulacan. Upon receiving the approved permit, the researcher coordinated with the principal of the school respondents through phone or messenger for the schedule of quantitative data collection. Since face-to-face classes are implemented this school year, the researcher collected the quantitative data personally.

There are two types of data that were collected, the quantitative and the qualitative data. In the quantitative data gathering, the questionnaire that was utilized is composed of two parts. Part I of the questionnaire provided information on the profile of the teacher respondents which includes their age, sex, educational attainment, specialization, and seminars attended related to test construction. Meanwhile, Part II of the questionnaire is composed of the 40-item Teachers' Competence Questionnaire (TECOM-Q) adapted from Hamafyelto, Hamman-Tukur, and Hamafyelto (2015) which were utilized to measure the teachers' test construction competencies. The scale was rated as 5 Always, 4 Almost Always, 3 Frequently, 2 Sometimes, and 1 Not At All.

Selected Science teachers were asked to complete Part I and II of the said survey form. Moreover, the respective department head of the selected teachers validated the teachers' perceptions by filling out only Part II of a similar survey questionnaire using teacher-made tests as basis. And, in cases where the school does not have a Science department head teacher, the school head was asked to validate the teachers' perceptions.

Following the conduct of survey is qualitative data gathering through a semi-structured interview. Interviews were conducted after written consent is provided by the respondents. Inclusions in the endorsement from the division office to conduct the study are non-disruption of regular instructional programs and ensuing of standard health safety protocols. In order to abide by these rules, qualitative data gathering was achieved through in-person interview for respondents from San Miguel National High School and Vedasto R. Santiago High School. Meanwhile, respondents from John J. Russell Memorial High School, Partida National High School, and Emilia Perez Ligon High School (formerly Maligaya High School) were interviewed through Messenger calls. Messenger call was used for the interviews to lessen the exposure of the teachers and students as stated in the endorsement of the Department of Education regarding health protocols. However, since the researcher is currently teaching in San Miguel National High School, she finds that respondents from her school and Vedasto R. Santiago High School are accessible for the in-person interview. She did this during her free time without interruption of any classes.

A written consent was essentially secured by the researcher to ensure that ethical considerations are followed, particularly on how the data will be recorded, evaluated, and stored. Personal information provided by respondents were treated with care and utmost confidentiality. The researcher ensured that the study is completed before all data gathered were properly disposed or deleted. The research was completed by the researcher on December 2022.

➤ *Sampling Procedures*

For the teacher and department head teacher respondents, total enumeration was applied. All teachers teaching Science in Grade 7 and 8 and their respective head teachers were requested to participate in the study.

However, for the student respondents, purposive sampling was utilized. Only the Grades 7 and 8 students were requested to participate in quantitative data gathering. Bernard (2002) claims that the judgment sampling method, also known as purposive sampling, is the purposeful selection of a participant based on the participant's personal characteristics. This nonrandom technique does not require underlying theories or a predetermined number of participants. Simply defined, the researcher chooses what information is necessary to have and then searches for sources willing and able to supply it based on their knowledge or experience. The researcher decided to use only Grades 7 and 8 since most Science subjects are offered in these grade levels.

Additionally, since the population (N=6834) is too large, first quarter grades of only ten percent of this population (n=683) were collected. For descriptive research the sample should be 10% of the population for a larger population as large as 1000. But if the population is small (as small as 500 and below) then 20% may be required. In correlational research, a relationship must be established between at least 30 subjects. The recommended minimum for experimental research is frequently 30 students per group.

The 683 student respondents were chosen using the lottery method sampling. In applying this method, the researcher picks numbers randomly, each number corresponding to the name of the student, in order to create the sample. The researcher made sure that the numbers are evenly distributed in order to build a sample in this manner.

Table 1 Respondents of the Study

School	Teachers	Department/School Head	Grade 7 and 8 Students	
			N	N
San Miguel National High School	21	1	3690	369
Vedasto R. Santiago National High School	6	1	1353	135
John J. Russell Memorial High School	6	1	1149	115
Partida National High School	3	1	513	51
Maligaya High School	2	1	129	13
Total	38	5	6834	683

For the qualitative data, two teachers and students per school were chosen to participate in the interview. This produced ten teachers and ten students for utilizing purposive sampling. Purposive sampling provided the researcher a leeway to decide what needs to be known and determines who among the sample can and are willing to provide the necessary data by virtue of knowledge and experience (Etikan, 2016). Lottery method sampling was utilized in choosing the respondents for interview. Respondents for the interview were given prior information about the topics that was discussed during the said interview for them to be familiar with the flow and context of the conversation and comfortably provide rich information about the variables under study.

➤ *Data Analysis Scheme*

Once all the questionnaires were collected, data were organized, tallied, tabulated, and analyzed using some statistical tools, which for the descriptive statistics employs range, mean and standard deviation to describe the pupils' academic performance in Science. Also, weighted mean was computed to describe the teacher respondents' test construction competencies. In addition, t-test for independent samples was used to determine if significant difference existed between the assessments of the school heads and the teachers themselves regarding their test construction competencies. Meanwhile, F-test and t-test were applied to determine if significant difference existed between and among the Science teachers' test construction competencies when they are grouped according to their demographic profile. Lastly, correlation analysis was performed to determine if significant relationship existed between the independent variables (teacher respondents' test construction competencies) and dependent variable (pupils' academic performance in Science).

For the gathered qualitative data, content analysis was used for interpretation. Content analysis is being used in a variety of research application particularly in the field of education as it seeks to cultivate and delve on the social reality or phenomenon through interpreting verbal recorded communication materials through identifying themes (Shava et.al, 2021).

### CHAPTER THREE RESULTS AND DISCUSSIONS

This chapter deals with the presentation, analysis, and interpretation of the data collected and the results of the statistical treatment employed in the study with the purpose of determining the relationship between the test construction competencies of public high school Science teachers and the academic performance of their respective students.

➤ *Socio-Demographic Profile of the Science Teachers*

A mix of social and demographic characteristics that characterize members of a given group or population are referred to as socio-demographics. In other words, when we discuss socio-demographics, we mean various social and demographic traits that enable us to understand what characteristics a group's members share.

Tables 2 to 6 show the socio-demographic profile of the Science teachers in terms of age, sex, highest educational attainment, specialization, and seminars attended related to test construction.

➤ *Age*

Age is measured by calculating the time elapsed (usually in complete years) between date of birth and a specific point in time (e.g. date of a particular survey).

The distribution of Science teacher respondents according to age is exhibited in Table 2.

Table 2 The Profile of the Science Teachers in terms of Age

Age	F	%
53 – 59	5	13.16
46 – 52	5	13.16
39 – 45	6	15.79
32 – 38	12	31.58
24 – 31	10	26.32
Total	38	100.00

The table shows that the most numerous age group is between 32 to 38 years old, which accounts for 31.58 percent of the respondents. The next largest age group, at 26.32 percent, consists of teachers aged 24 to 31, 15.79 percent of those aged 39 to 45, and 13.16 percent of those aged 46 to 52 and 53 to 59.

The results imply that most of the Science teachers are in the middle-aged group and just starting their career in the teaching profession.

➤ *Sex*

Sex refers to a set of biological attributes in humans. It is primarily associated with physical and physiological features including chromosomes, gene expression, hormone levels and function, and reproductive/sexual anatomy.

Table 3 The Profile of the Science Teachers in terms of Sex

Sex	F	%
Male	8	21.05
Female	30	78.95
Total	38	100.00

The distribution of science teacher respondents according to sex is exhibited in Table 3.

Table 3 shows that the majority or 78.95 percent of the Science teachers are female. Meanwhile, the remaining 21.05 percent of this group of teachers are male.

The results imply that females dominated the Science teachers in public secondary schools in San Miguel, Bulacan.

According to Bongco and Ancho (2020), teaching is a profession that has been dominated by women. This is supported by the data from The World Bank that shows that as of 2020, 71% of secondary school teachers are female (The World Bank, retrieved December 2022.) Underrepresentation of males in the teaching world is a pressing matter. Interestingly, The World Bank data shows that the percentage of females teaching in public secondary schools is decreasing steadily. This gives hope that there will be a balance in the number of female and male teachers in the future.

➤ *Highest Educational Attainment*

An individual's educational achievement is the highest level of schooling they have successfully completed. The attainment of learning objectives for a level of education is referred to as its successful completion. These learning objectives are often validated through the assessment of gained information, skills, and competences.

Table 4 The Profile of the Science Teachers in terms of Highest Educational Attainment

Highest Educational Attainment	F	%
Doctorate Degree	2	5.26
Master's Degree	19	50.00
Bachelor's Degree	17	44.74
Total	38	100.00

The distribution of science teachers when they are classified according to the highest level of educational attainment is exhibited in Table 4.

Reflected from the table that exactly one-half, or 50 percent of the Science teachers have master's degree. Meanwhile, 44.74 percent have bachelor's degree, and the remaining 5.26 percent have doctorate degree.

The results imply that more Science teachers have attained higher education, but some teachers are not pursuing graduate programs. This may be caused by lack of time due to lot of extra assignments in their schools. Teachers have a lot of work especially those who serve as advisers are required to participate in government programs such as feeding programs, deworming, population census, anti-drug campaigns, elections, mass immunization community mapping, conditional cash transfer (4 P's), etc. They are also required to monitor their students, especially those who have a habit of incurring absences and those who are in danger of dropping out.

In relation to this, Esguerra (2018) stated that it is well-known that teachers in public schools are overworked and overloaded. Teachers in the Philippines are not only teaching but also have non-teaching tasks. Multiple tasks are given to each teacher, thus decreasing teachers' time in pursuing higher education. (David, Albert and Vizmanos, 2019.)

➤ *Specialization*

In academics, specialization is a course of study or major field in which a specialist practices. In education, academic specialization pertains to the subject that they specialize in and teach.

Table 5 The Profile of the Science Teachers in terms of Specialization

Specialization	F	%
General Science	8	21.05
Biology	18	47.37
Chemistry	3	7.89
Physics	5	13.16
Social Science	1	2.63
TLE	1	2.63
others	2	5.26
Total	38	100.00

The distribution of Science teachers when they are grouped according to their demographic profile in terms of specialization is indicated in Table 4.

According to the data, 47.37% have specialized in. 21.05 percent in General Science, 13.16 percent in Physics, and 7.89 percent in Chemistry. Non-science majors include 2.63 in Social Science and TLE and 5.26 in others.

The results imply that there are some teachers who teach Science but are not major of in Science. Out of 38 Science teacher respondents, 4 of them have specialized in other subjects (T.L.E., Social Science and others.)

According to Hobbs (2015), many teachers are teaching subjects outside their field of expertise. This has been a practice that many have experienced. Teacher Education Programs ensure the quality of school teachers they produce, although once in the field of teaching, teachers may or may not be assigned to teach their area of specialization. One reason includes the lack of teachers specializing in other subjects. Although out-of-field teachers have the probability of producing low levels of student achievement, teachers have almost always been flexible as regards teaching strategies. They can adapt well to cater to the needs of their learners.

➤ *Seminars Attended Related to Test Construction*

A seminar is a format of instruction that occur in several meetings. Universities, professional organizations, or commercial associations generally offer seminars to provide academic instruction in a small group setting. In education, seminars on test construction give a combination of lectures, group activities and assessments to form adequate understanding on the subject.

Table 6 The Profile of the Science Teachers in terms of Seminars Attended Related to Test Construction

Seminars	F	%
Yes	25	65.79
No	13	34.21
Total	38	100.00

It can be gleaned from the table that the majority or 65.79 percent of the Science teachers claimed that they already attended seminars related to test construction. On the other hand, more than one-third or 34.21 percent stated that they did not attend any seminar on test construction.

The results imply that there are some Science teachers who lack seminars in test construction.

Based on the study conducted by Ahmed (2018), test construction skills training proved to be significant in developing and improving the test construction skills of teachers. In-service teachers need to be consistently trained as regards test construction. Ahmed (2018) recommended that seminars should include test construction improvement training. Lack of seminars and training conducted as regards test construction is detrimental to the quality of tests that will be administered to the students.

➤ *The Test Construction Competencies of the Public High School Science Teachers*

The development of assessment instruments is necessary for assessing students' mastery of course content and level of application skills. A well-designed assessment will result in more accurate outcomes. One strategy for ensuring that students learn from their assessments is for teachers to create them properly. In addition, it offers teachers enhanced data for instruction. The primary objective of the evaluation procedure is to collect valid, dependable, and valuable information about the students.

The test construction competencies of the public high school Science teachers in terms of developing assessment method appropriate for instructional decisions; administering, scoring, and interpreting results; using assessment results; developing valid students grading procedures; communicating assessment result; and recognizing unethical and illegal assessment methods are presented in Tables 7 to 12.

➤ *Developing Assessment Method Appropriate For Instructional Decisions*

Teachers should be able to use assessment data effectively when making decisions about individual students, organizing instruction, developing curricula, and promoting school development. They should be able to develop valid grading systems based on student evaluations.

Table 7 The Test Construction Competencies of the Public High School Science Teachers in terms of Developing Assessment Method Appropriate for Instructional Decisions

Item Statement The Science Teacher...	School Head		Teacher	
	Mean	VD	Mean	VD
1. Understands How Valid Assessment Data Supports Instructional Activities.	3.86	Aa	4.21	A
2. Diagnoses Group And Individual Needs Of Student.	4.22	A	4.26	A
3. Uses Test Results To Motivate Students.	4.28	A	4.48	A
4. Uses Test Results To Evaluate Instruction.	4.22	A	4.39	A
5. Uses Evaluation Assessment Options In Giving Tests.	4.18	Aa	4.26	A
6. Considers The Cultural Background Of Students In Test Construction.	4.16	Aa	4.37	A
7. Sees Social Background Of Students As An Important Factor In Test Construction.	3.68	Aa	4.28	A
8. Considers Economic Background Of Students As An Important Factor In Test Construction.	3.72	Aa	4.18	Aa
9. Uses Test Results To Make Appropriate Decision About The Students	4.22	A	4.34	A
10. Collects Information That Facilitates One's Decision About A Test	3.88	Aa	4.21	A
11. Uses Students' Test Data To Analyze The Quality Of One's Assessment Techniques.	3.87	Aa	4.23	A
12. Uses Various Assessment Methods In Teaching One's Subject.	4.32	A	4.66	A
13. Avoids Common Mistakes In Student Assessment.	4.22	A	4.23	A
Overall Mean	4.08	Aa	4.32	A

➤ *Legend:*

Scale	Verbal Description
4.21 – 5.00	Always (A)
3.41 – 4.20	Almost Always (AA)
2.61 – 3.40	Frequently (F)
1.81 – 2.60	Sometimes (S)
1.00 – 1.80	Not at All (N)

Table 7 displays the test construction competencies of Public High School Science teachers in terms of developing assessment Method Appropriate for Instructional Decisions.

Drawn from the table that out of the 13 items indicated therein, the school heads gave their teachers a verbal description of “almost always” in five items and “always” in six items. Meanwhile, teachers rated themselves with a verbal description of “almost” in only one item and “always” in the remaining 12 items.

The results imply that both the school heads and teachers agree that these teachers always diagnose group and individual needs of students, use test results to motivate students, use test results to evaluate instruction, use test results to make appropriate decisions about the students, use various assessment methods in teaching one’s subject, and avoid common mistakes in student assessment. On the other hand, teachers had a greater assessment of themselves as compared to the assessment of their school heads in terms of understanding how valid assessment data support instructional activities, using evaluation assessment options in giving tests, considering the cultural background of students as an important factor in test construction, collecting information that facilitates one’s decision about a test, and using students’ data to analyze the quality of one’s assessment techniques.

A closer look at the table shows that all the items considered, the school heads assessed that their teachers are “almost always” competent as regards test construction in terms of developing assessment method appropriate for instructional decisions. On the other hand, the teachers themselves believe that they are “always” competent as regards this variable.

The results imply that school heads have a lower assessment of their teachers compared to the assessment of the teachers themselves as regards these teachers’ test construction competencies in terms of developing assessment methods appropriate for instructional decisions.

All items in table 7 indicated therein have a computed overall mean of 4.08 for school heads rating the teachers with a verbal description of “almost always” and 4.32 for teachers rating themselves with a verbal description of “always.” Notice the item “considers the economic background of students as an important factor in test construction” had the lowest weighted mean, with school heads rating the teachers 3.72 and teachers rating themselves 4.18 with a verbal description of “almost always.” The highest weighted mean was for the item “uses various assessment methods in teaching one’s subject,” wherein school heads rated teachers a 4.32 and teachers rated themselves a 4.66, both with the verbal description of “always.”

The results imply that both school heads and teachers agree that these teachers are keen to establish assessment methods that are appropriate for teaching. It has been observed that they take care to consider the appropriate assessment tools provided to students as they prepare for them, and that they do so with appropriate considerations and assessment administration to ensure student learning. In addition, they determine how well they function, particularly how well teachers can handle a data-driven assessment analysis and how much teachers can learn about it from the answers or responses students provide.

In contrast to the results of the present study, Lam (2015) asserted that teachers still need to be fully proficient in the design and implementation of assessments. Even though it has been said that assessment should be marked to assist students in learning, most teachers need to learn how to use formative assessment in the classroom or summative assessment information to improve teaching and learning. The study of Lam (2015) investigates the landscape of language assessment training in five Hong Kong teacher education institutions against the backdrop of assessment reforms in primary/secondary school settings. Specifically, he examined the extent to which two assessment courses may help or impede the development of language assessment literacy among pre-service teachers at a single-teacher education school. His study determined that language assessment training in Hong Kong still needs to be improved. Specific language assessment courses need to be revised to bridge the theory-practice gap within the framework of assessment reform.

In the interview, teacher respondents were asked how they guarantee and verify that their teaching and instructional decisions align with assessment procedures. Most teachers responded that they claim and offer proof that all their decisions are consistent with assessment procedures. In addition, they mentioned that before teaching, they must carefully plan, utilizing the curriculum map and learning plan, to ensure everything is in order.

In the conducted interview with the students, they were asked to describe the tests made and administered by their science teacher. Some students replied that their tests are a mix of easy, tricky, and hard questions. Some said that the questions are mostly easy. A few of the students added that most of the time their teacher gives questions with no choices, which they find more difficult to answer than those with choices given.

➤ *Administering, Scoring and Interpreting Results*

Assessment is an essential component of education since it analyzes whether educational goals are being met. It influences decisions regarding grades, placement, progression, instructional requirements, curriculum, and in some cases, funding. Scoring entails assigning numeric values to an individual's performance on a certain task. The analysis and interpretation of data should support the central purpose of assessment, which is to determine and appreciate where learners stand concerning a specific learning area at the time of assessment.

Table 8 The Test Construction Competencies of the Public High School Science Teachers in terms of Administering, Scoring and Interpreting Results

Item Statement The Science Teacher...	School Head		Teacher	
	Mean	VD	Mean	VD
1. Assesses Students' Performance In Class Assignments.	4.32	A	4.36	A
2. Assesses Students' Performance In Homework Assignment.	4.36	Aa	4.34	A
3. Uses Guide For Scoring Essay Type Questions And Projects.	4.72	A	4.88	A
4. Uses Stencils For Scoring Response Choice Questions.	3.32	F	4.08	Aa
5. Uses Scales For Rating Performance Assessment.	4.36	A	4.29	A
6. Can Interpret Report Scores In Percentile.	4.52	A	4.68	A
7. Can Correct Raw Scores Into Standardized Scores.	4.18	Aa	4.20	Aa
Overall Mean	4.25	A	4.40	A

➤ *Legend:*

Scale	Verbal Description
4.21 – 5.00	Always (A)
3.41 – 4.20	Almost Always (AA)
2.61 – 3.40	Frequently (F)
1.81 – 2.60	Sometimes (S)
1.00 – 1.80	Not at All (N)

The test construction competencies of the Public High School Science teachers in terms of administering, scoring, and interpreting results are presented in Table 8.

Out of the seven items indicated above, the school heads gave their teachers a verbal description of “always” in four items, “almost always” in two items, and “frequently” in one item. Meanwhile, the teachers rated themselves with a verbal description of “always” in five items and “almost always” in two items.

The results imply that both the school heads and the teachers are in congruence that the Science teachers are “always” assessing students' performance in class assignments, using guides for scoring essay-type questions and projects, using scales for rating performance assessment, and can interpret scores in percentile. Notably, both the school heads and the teachers acknowledge that these teachers are “almost always” able to correct raw scores into standardized scores. On the other hand, school heads only “almost always” believed that their teachers are assessing students' performance in homework exams, wherein teachers believed that they “always” do so. Additionally, school heads assessed that their science teachers “frequently” use stencils for scoring response choice questions, in which the teachers stated that they “almost always” practice doing it.

The overall mean of all the items indicated that both the school heads and the teachers “always” believed that these teachers have a high level of test construction competencies in terms of administering, scoring, and interpreting test results.

Drawn from all the items in Table 8, the computed overall mean is 4.25 for school heads' rating of their teachers and a mean of 4.40 for the teachers' rating of themselves, garnering both a verbal description of "always." It appears that the item "uses stencils for scoring response choice questions" had the lowest weighted mean, with school heads' rating their teachers a mean of 3.32 with a verbal description of “frequently” and teachers' rating of themselves with a mean of 4.08 with a verbal description of "almost always." The highest weighted mean was for the item " uses guide for scoring essay type questions and projects," wherein teachers were rated by their school heads a mean of 4.72 and teachers rated themselves a mean of 4.88, both with the verbal description "always."



The results imply that these two groups of respondents recognize that science teachers administer, score, and interpret the assessment outcomes. They adopt the proper assessment procedures to ensure academic integrity and honesty. Specifically, teachers play a significant role since they must properly verify that students' assessments are legitimate and dependable. They also ensure that their inputs, particularly the grading, is accurate and error-free. Thus, this is essential for students, as it will serve as the basis for their academic performance.

Consistently with the present study, Grahams et al. (2015) did a meta-analysis of real and quasi-experiments conducted with children in grades 1 to 8 to investigate whether formative writing evaluations directly related to ordinary classroom teaching and learning improve students' writing performance. They discovered that feedback on students' writing from adults, peers, themselves, and computers statistically improved writing quality, with effect sizes of 0.87, 0.58, 0.62, and 0.38, respectively. However, teachers' monitoring of students' writing development and adoption of the 6 + 1 Trait Writing model did not significantly improve students' writing. The results of this meta-analysis support the use of formative writing exams that provide direct feedback to students as an integral element of everyday teaching and learning. Therefore, they believe that teachers should utilize such tests more frequently and that they should play a more significant part in the Next-Generation Assessment Systems being developed by Smarter Balanced and PARCC.

In the interview, respondents were asked how they ensure proper assessment processes are adhered to when administering, scoring, and interpreting results. According to many teachers, the assessment was easier to administer because everything had been meticulously arranged. Other teachers have stated that they have a method that must be adhered to; therefore, they can guarantee that the assessment process was conducted correctly. In addition, teachers stated that they are to emphasize the importance of ensuring that students understand the lesson's objectives before implementing them. In addition, teachers reported that they had offered a tool for scoring, calculating the number of suitable responses, and analyzing results.

During the interview with the students, all students commented that their teachers give tests about topics that were taken in the past lessons. In the follow-up question, many of the students reply that the instructions and questions given by their science teacher were very clear and easy to understand, thus helping them to answer correctly. However, a few of the student respondents said that sometimes there are questions that are a little bit harder to understand at first, so the teacher will explain the instructions further.

#### ➤ *Using Assessment Results*

Assessment results are essential evidence that can be used to support funding requests, make changes to the curriculum, rethink faculty lines, and more. Positive results on an assessment are a good sign of learning. However, disappointing (negative) test results can also be helpful if used to improve the learning process.

Table 9 The Test Construction Competencies of the Public High School Science Teachers in terms of Using Assessment Results

Item Statement The Science Teacher...	School Head		Teacher	
	Mean	VD	Mean	VD
1. Interprets Accumulated Assessment Information.	4.22	A	4.36	A
2. Uses Accumulated Assessment Information To Facilitate Students Learning.	4.38	A	4.52	A
3. Uses Assessment Results To Evaluate Teaching Methods.	4.28	A	4.38	A
4. Interprets Test Results To Interpret Students' Learning.	4.88	A	4.92	A
5. Uses Test Results As Guide For Students' Remediation.	4.72	A	4.88	A
Overall Mean	4.49	A	4.61	A

#### ➤ *Legend:*

Scale	Verbal Description
4.21 – 5.00	Always (A)
3.41 – 4.20	Almost Always (AA)
2.61 – 3.40	Frequently (F)
1.81 – 2.60	Sometimes (S)
1.00 – 1.80	Not at All (N)

The test construction competencies of the Public High School Science teachers in terms of using assessment results are presented in Table 9.

Among the five items indicated therein, the school heads and the teachers agree that the teachers “always” practice the given variables. It is indicated by the five ratings of “always from the school heads’ assessment of their teachers and five ratings of “always” from the teachers’ assessment of themselves.

Further observation of the tabulated findings reveals that although the school heads and the teachers recognize that the teachers “always” practice these variables, the school heads’ assessment of the teachers are lower in all variables as compared to the teachers’ assessment of themselves. This is also true as regards the overall mean for these items.

The results imply that the teachers have a higher assessment of themselves as regards their test construction skills in terms of using assessment results as compared to their school heads’ assessment of them.

The item statements have the overall computed mean as school heads rated teachers with 4.49 and teachers rated themselves with 4.61 both with a verbal description of "always." The item "interprets accumulated assessment information" had the lowest weighted mean, with school heads' rating their teachers a mean of 4.22 and teachers rating themselves a mean of 4.36 with a verbal description of "always." Meanwhile, the highest weighted mean was for the item "interprets test results to interpret students' learning," wherein school heads rated teachers a 4.88 and teachers rated themselves a 4.92, both with the verbal description "always."

The results imply that teachers, as validated by their school heads, view assessment results as key evidence for the overall curriculum and student learning. It would provide teachers with a basis for evaluating student learning. They can adapt based on the outcomes to ensure that their students are learning continually and effectively. The proper use of assessment will show teachers how important it is for students to learn.

Based on Brabeck et al. (2016), the study is congruent with the present study's findings. They argued that teacher preparation programs have both a desire and a responsibility to prove that teacher education affects Pre-K–12 student learning. To make decisions regarding students' progress, whom to suggest for state licensure, and how to improve teacher education, program faculty require accurate data. Consequently, it offers suggestions for teacher educators to improve whoever wishes to enhance their professional programs to continue improving students' learning.

During the interview, respondents were asked how they assure assessment data is considered when assessing instructional approaches and student growth. The teachers ensure that the data is effectively utilized for their advantage and the learning of the students. Teachers said that it is essential to consider assessment data for evaluating student growth and teaching tactics. Other teachers stated that it will provide them with everything they need, assist them in developing instructional techniques that are tailored to the needs of the students, and speed up receiving assessment results, all of which are believed to contribute to the success of the students because of solid teaching practices. On the other hand, some responded that they would utilize it to assure that students are learning through effective instruction derived from the outcomes of student evaluations. According to them, it is crucial to monitor the quality of curriculum development and implementation based on the success or failure of assessment outcomes.

#### ➤ *Developing Valid Students Grading Procedures*

Grading is the process through which a teacher assesses student learning through classroom tests and tasks, the framework in which competent teachers build this process, and the conversation that surrounds grades and determines their meaning for possible interpretation. The objective of a grading system is to provide students with feedback so they can assume responsibility for their education and to provide information to those who support these learners (teachers, school heads, parents, and others).

Table 10 The Test Construction Competencies of the Public High School Science Teachers in terms of Developing Valid Students Grading Procedures

Item Statement The Science Teacher...	School Head		Teacher	
	Mean	VD	Mean	VD
1. Uses Grades As Punishment For Erring Students.	2.24	S	2.12	S
2. Modifies Grading Procedures At Different Times.	4.18	Aa	4.20	Aa
3. Ensures That The Grading System He/She Uses Follows The Standard Grading System Issued By The Department Of Education.	4.88	A	4.94	A
4. Sees To It That The Students Understand The Grading System.	4.74	A	4.88	A
5. Ensures That Parents Are Familiar With The Grading System.	4.72	A	4.90	A
Overall Mean	4.15	Aa	4.21	A

#### ➤ *Legend:*

Scale	Verbal Description
4.21 – 5.00	Always (A)
3.41 – 4.20	Almost Always (AA)
2.61 – 3.40	Frequently (F)
1.81 – 2.60	Sometimes (S)
1.00 – 1.80	Not at All (N)

Table 10 displays the test construction competencies of public High School Science teachers in terms of developing valid student grading procedures.

Out of the five items specified therein, the school heads and the teachers rated these teachers with a verbal description of “always” in three items, “almost always” in one item, and “sometimes” in the remaining item.

The results imply that these two groups of respondents agree that the teachers only “sometimes” use grades as punishment for erring students. They are also in congruence that the science teachers “almost always” modify grading procedures at different times. Notably, they both concurred that the teachers “always” ensure that the grading system he/she uses follows the rules issued by the Department of Education, and that the parents are aware of the grading system.

Interestingly, even though these two groups of respondents both rated the teachers with similar verbal descriptions for all the indicated items, the verbal description for the overall mean for the teachers’ test construction competencies in terms of developing valid student grading procedures came out to be different. This is indicated by the lower value of ratings given by the school heads to their teachers as compared to the teachers’ assessment of themselves. Overall, the school heads gave their teachers a rating of “almost always”, meanwhile the teachers “always” believed that they are competent as regards test construction in terms of the aforementioned variable.

Analysis of Table 10 has a computed overall mean of 4.15 as school heads rated their teachers with a verbal description of “almost always” and 4.21 for teachers rating themselves with a verbal description of “always.” Assessing all the items, the lowest weighted mean was incurred by the item “uses grades as punishment for erring students,” wherein school heads rated teachers a mean of 2.24 and teachers rated themselves with a mean of 2.12, both with a verbal description of “sometimes.” It is notable that the item “ensures that the grading system he/she uses follows the standard grading system issued by the Department of Education” had the highest weighted mean, with school heads rating teachers a mean of 4.88 and teachers rated themselves a mean of 4.94 with a verbal description of “always.”

The results imply that teachers, as agreed by their school heads, ensure that the proper grading procedure is adhered to. It has been observed that they appreciate the significance of having a clear criterion for assessing students’ learning that is standardized within their own school and across the country utilizing the grading system approved by the Department of Education. Proper communication is also essential to this method, particularly with the teachers, who must convey this information to the students and parents so that they may interpret the assessment findings based on the grading procedure. It is worth noting that they recognize that grades are not meant to punish but rather to assess learning and serve as a reminder of whether one should exert more effort, as seen by their performance in school as indicated by these grades.

In connection to the present study, Cook et al. (2015) noted that the ultimate goal of all assessments is to facilitate a defensible conclusion regarding the assessed individual. Validation is the procedure of gathering and analyzing information supporting a decision. The rigorous validation process entails articulating the claims and assumptions connected with the proposed decision (the interpretation/use argument), experimentally evaluating these assumptions, and organizing evidence into a cohesive discussion of validity. They supported the idea of Kane where he specifies four inferences in the validity argument: Scoring (converting an observation into one or more scores), Generalization (using the scores as a reflection of performance in a test context), Extrapolation (using the scores as a reflection of actual performance), and Implications (applying the scores to inform a decision or action). Each of these conclusions should be supported by evidence, emphasizing the most dubious assumptions in the chain of reasoning. Depending on the assessment’s intended application or associated choice, the key hypotheses (and required proof) differ. Thus, the Kane framework applies to quantitative and qualitative assessments, individual exams, and assessment programs.

In the interview, respondents were asked how they ensure the validity and reliability of the grading procedure. The teachers were asked how they perceived the validity and dependability of the method. Many teachers said that the legitimacy and dependability of the grading system were ensured because the assessment methods aligned with the instructional strategies. According to some teachers, they must rigorously evaluate and examine test data to build a transparent and trustworthy grading system. Additionally, some teachers responded that they are prepared to steer themselves toward a valid and dependable grading system. Other teachers said that they checked the alignment of student grades with the grading system very carefully.

During the interview with the students, when asked, “How clear are the directions and questions in the tests given by your science teacher?” Some students shared that their teachers are really keen on explaining the directions given on the tests so that they may easily understand the questions. Few of them said that their teachers give directions and expects everyone to understand it right away, although some of their classmates sometimes needed to ask clarifications.

In the follow-up question, students were asked, “How competent is your science teacher in formulating questions in exams?” Many students said that they’re teachers are very competent in formulating questions. They shared that although the questions were sometimes hard, it is understandable because the subject is also hard. The questions were easy to understand, and the

instructions were always clear. On the other hand, some students claimed that their teachers are not very competent in formulating questions because they sometimes just copy the questions already present in the modules and books.

➤ *Communicating Assessment Results*

The fundamental objective of presenting assessment results should always be to motivate action, particularly among students and parents. In this way, results have the greatest chance of being utilized if they convey a relevant story, clear, concise, and convincing, and provide adequate responses to legitimate critiques.

Table 11 The Test Construction Competencies of the Public High School Science Teachers in terms of Communicating Assessment Results

Item Statement The Science Teacher...	School Head		Teacher	
	Mean	VD	Mean	VD
1. Discusses Assessment Results With Other Teachers.	4.22	A	4.24	A
2. Uses Appropriate Terminology In Reporting The Meaning Of Assessment.	4.58	A	4.72	A
3. Explains The Importance Of Taking Examinations To His/Her Students.	4.74	A	4.82	A
4. Recognizes Measurement Errors In His/Her Assessment.	3.86	Aa	4.18	Aa
5. Understands The Limitation Of The Assessment Method.	4.26	A	4.24	A
Overall Mean	4.33	A	4.44	A

➤ *Legend:*

Scale	Verbal Description
4.21 – 5.00	Always (A)
3.41 – 4.20	Almost Always (AA)
2.61 – 3.40	Frequently (F)
1.81 – 2.60	Sometimes (S)
1.00 – 1.80	Not at All (N)

The test construction competencies of the Public High School Science teachers in terms of communicating assessment results are presented in Table 11.

Among the five items indicated therein, both the teachers and their school heads assessed that the science teachers “almost always” in one item, and “always” in four items. The overall mean also had the same verbal description of “always” for both groups.

The results imply that as validated by their school heads, science teachers “almost always” recognize measurement errors in his/her assessment. Notably, they both agree that science teachers discuss assessment results with other teachers, use appropriate terminology in reporting the meaning of assessment, explain the importance of taking examinations to his/her students and understand the limitation of the assessment method.

Analysis of Table 11 has a computed overall mean of 4.33 for school heads’ rating of their teachers and of 4.44 for the teachers’ rating themselves both with a verbal description of “always.” Assessing all the items, the lowest weighted mean was for the item “recognizes measurement errors in his/her assessment,” wherein school heads rated teachers 3.86 and teachers rated themselves 4.18, both with a verbal description of “almost always.” The item “explains the importance of taking examinations to his/her students” had the highest weighted mean, with school heads rating their teachers with a mean of 4.74 and teachers rating themselves with a mean of 4.82 with a verbal description of “always.”

These results imply that the science teachers, as validated by their school heads, effectively convey assessment results to students. They agree that evaluation is meaningful if effectively communicated to students, particularly intending to incorporate assessment into the learning process. Notably, teachers demonstrated an excellent environment for assessment, allowing students to recognize its significance for their learning. In addition, teachers are viewed as an integral component of the communication process, particularly when addressing the questions of people who wish to comprehend the information provided, especially those parents who seek clarification. From this information, assessment communication is vital in all school settings.

On the contrary, Sadler noted that even with good-quality assessment feedback from the teacher, students show little or no development at all. He stated that it is a common and bemusing observation that improvement in students’ work, tests, and projects does not imperatively follow even after giving meaningful, reliable, and valid feedback. (Sadler 1989 as cited in Tran N. 2014.)

In the interview, respondents were asked how they guarantee that the intended recipient's evaluation results are successfully communicated. Teachers were asked how they ensure that information reaches the target receptor effectively. Many teachers responded that they were only facilitators and graders. Other teachers mentioned that they must maintain the privacy of their students' exams and grade outcomes because they know their implications. Teachers also said they know how to get in touch with and talk to parents to ensure they get all results.

➤ *Recognizing Unethical And Illegal Assessment Methods*

Recognizing these malpractices, particularly concerning pedagogical approaches, is a proactive indicator that one might avoid slipping into these. It is the responsibility of teachers to be able to recognize whether a method of assessing students or the use of assessment results violates ethical standards, is illegal, or is inappropriate in some other way.

Table 12 The Test Construction Competencies of the Public High School Science Teachers in terms of Recognizing Unethical and Illegal Assessment Methods

Item Statement The Science Teacher...	School Head		Teacher	
	Mean	VD	Mean	VD
1. Is Familiar With The Laws Prohibiting Exam Malpractice.	3.22	F	3.88	Aa
2. Is Familiar With Procedures That Lead To Misuse Of Assessment Results.	3.88	Aa	4.20	Aa
3. Is Familiar With Procedures That Lead To Overuse Of Assessment Results.	3.76	Aa	4.16	Aa
4. Is Familiar With The Students' Rights To Confidentiality In Assessment.	4.28	A	4.19	Aa
5. Considers Tests As A Measure Of His/Her Teaching Effectiveness.	4.52	A	4.40	A
Overall Mean	3.93	Aa	4.17	Aa

➤ *Legend:*

Scale	Verbal Description
4.21 – 5.00	Always (A)
3.41 – 4.20	Almost Always (AA)
2.61 – 3.40	Frequently (F)
1.81 – 2.60	Sometimes (S)
1.00 – 1.80	Not at All (N)

Table 12 displays the test construction competencies of Public High School Science teachers in terms of recognizing unethical and illegal assessment methods.

Seen from the table that the school heads rated their teachers with a verbal description of “always” in two items, “almost always” also in two items, and “frequently” in one item. Meanwhile, the teachers rated themselves with a verbal description of “almost always” in four items, and “always” in one item. Both groups got the same overall verbal description of “almost always” as regards the science teachers’ test construction competencies in terms of recognizing unethical and illegal assessment methods.

The results imply that both these groups of respondents agree that the science teachers “almost always” are familiar with procedures that lead to misuse of assessment results and with the procedures that lead to overuse of assessment results. Additionally, they both acknowledge that science teachers “always” consider tests as a measure of his/her teaching effectiveness. On the other hand, school heads had a lower level of assessment of their teachers as compared to the teachers’ assessment of themselves in terms of these teachers being familiar with the laws prohibiting exam malpractice. School heads rated their teachers with a verbal description of “frequently” as regards this item. Meanwhile, the teachers said that they are “almost always” familiar with the aforementioned variable.

Interestingly, school heads had greater assessments of their teachers as compared to the teachers’ assessment of themselves in terms of being familiar with the students’ right to confidentiality in assessment. Teachers only “almost always” believed that they are competent as regards this item. But their school heads, on the contrary, “always” believed that their teachers are familiar with the students’ right to confidentiality in assessment.

Analysis of Table 12 reveals an overall mean of 3.93 for school heads’ rating of their teachers and 4.17 for teachers’ rating of themselves with the verbal descriptor “almost always.” The lowest weighted mean was for the item “is familiar with the laws prohibiting exam malpractice,” for which school heads gave teachers a 3.22 with a verbal description of “frequently” and teachers rated themselves a 3.88 with a verbal description of “almost always.” The item “considers tests as a measure of his/her teaching effectiveness” is the highest weighted means, with school heads rating teachers a mean of 4.52 and teachers’ rated themselves a mean of 4.40, both with the verbal description “always.”

The results imply that teachers are aware of the potential for unethical and unlawful assessment procedures, which indicates that they are working towards ethical and lawful assessment procedures. It is vital to notice that they know the entire process and the factors that could result in a mistake. Making sure this happens will result in a meaningful and reliable assessment for the

students. However, the effort required to search for such would be inconspicuous because most of the time, one learns about it after it has already occurred. Through familiarization, everything will be brought to hand and more effectively dealt with to present the students with the most efficient assessment possible.

In conjunction with the study, Shallianah (2016) perceived that it indeed needs to craft in determine and address issues concerning student assessments through a “Standards for Teacher Competence in Educational Assessment of Students.” The standards include teachers being skilled in choosing and developing assessment methods appropriate for instructional decisions; in administering, scoring and interpreting the results of various assessment methods; in using assessment results when making decisions about individual students, planning to teach, developing curriculum, and school improvement; in developing valid student grading procedures; in communicating assessment results to students, parents, co-educators, and other stakeholders; and in recognizing unethical, illegal, and inappropriate assessment methods and uses of assessment information.

During the interview, respondents were asked how vital it is to be aware of unethical and illegal evaluation processes. The teachers were asked how they viewed the significance of discovering potential evaluation malpractices. A significant number of teachers responded that it is vital to be aware of unethical and illegal evaluation methodological practices. In addition, teachers noted that it is their responsibility to serve as role models and act professionally.

➤ *The High School Students' Academic Performance In Science*

In this part of the study, the academic performance of Junior high school students in science is presented in Table 13.

Table 13 Distribution of Respondents when Classified According to Academic Performance in Science

Grade	F (N=683)	Percent	Verbal Description
90 and above	150	21.96	Outstanding (O)
85 – 89	149	21.82	Very Satisfactory (VS)
80 – 84	293	42.90	Satisfactory (S)
75 – 79	91	13.32	Fairly Satisfactory (FS)
74 and below	0	0.00	Did Not Meet Expectations (DNE)
Range	75 – 98		
Mean	84.31		
Verbal Description	Satisfactory		
Standard Deviation	6.22		

Table 13 shows that most of the students or 42.90 percent got grades between 80 and 84 with a verbal description of "satisfactory." Meanwhile, more than one-fifth or 21.96 percent obtained grades that lie within the bracket of 90 and above with a verbal description of "outstanding." A considerable portion, 21.82 percent obtained grades from 85 to 89 with a verbal description of "very satisfactory." The remaining 13.32 percent of the students got grades between 75 to 79 with a verbal description of "fairly satisfactory." Interestingly, no one incurred a grade of 74 or below, and a verbal description of "did not meet expectations".

Whereas the grades of the students ranged from 75 to 98 with a mean of 84.31 (satisfactory) and a standard deviation of 6.22. The results indicate that 464 students received grades from 78 to 91. This discloses that the grades of the students are heterogeneous in nature.

These results imply that despite the problems and obstacles that the students faced and experienced in this new normal, they were able to do well in science. This indicates that students have already adapted to their science teacher's way of designing and constructing assessment exams.

During the interview, students were asked to describe their academic performance in Science during the first quarter of this school year. Many of the students responded that they were satisfied with their academic performance and the grades they received. Some students said that they received higher grades than they expected. Few of them added that despite receiving satisfactory grades, they would strive to do better in the next quarter.

➤ *Test of Significant Difference Between the Assessments of School Heads and Teachers*

In this part of the study, Table 14 presents the results of the t-test analysis which was performed to determine if a significant difference existed between the assessments of the school heads and the Science teachers themselves as regards their test construction competencies.

Table 14 Results of the t-test Analysis on the Difference between the Assessments of School Heads and Teachers

Test Construction Competencies	Mean		Mean Diff.	T-Value	P-Value
	School Heads	Teachers			
Developing Assessment Method Appropriate For Instructional Decisions	4.08	4.32	-0.24	-3.496**	0.002
Administering, Scoring And Interpreting Results	4.25	4.40	-0.15	-0.754ns	0.465
Using Assessment Results	4.50	4.61	-0.11	-0.655ns	0.531
Developing Valid Student Grading Procedures	4.15	4.21	-0.06	-0.077ns	0.941
Communicating Assessment Results	4.21	4.33	-0.11	-0.527	0.612
Recognizing Unethical And Illegal Assessment Methods	3.93	4.17	-0.24	-0.978	0.357

Legend: \*\* = Significant ( $p \leq 0.01$ ) ns = Not Significant ( $p > 0.05$ )

Reflected from the table that a highly significant difference was found between the assessments of the school heads and the Science teachers themselves with regard to their test construction competencies in terms of developing assessment methods appropriate for instructional decisions. This highly significant difference is indicated by the computed probability value of 0.002 which is lower than the 0.01 significance level.

This implies that Science teacher respondents had greater assessments of themselves as compared to the assessment of their school heads in so far as test construction competencies in terms of developing assessment methods appropriate for instructional decisions. This may be because school heads did not have enough awareness of the teachers' knowledge on how to prepare tests in conjunction with their strategies used in teaching Science.

Further observation of the tabulated findings reveals that no significant difference was found between the ratings given by school heads to their teachers and the teachers' rating of themselves regarding their competencies in test construction in terms of administering, scoring, and interpreting results; using assessment results; developing valid students grading procedures; communicating assessment results; and recognizing unethical and illegal assessment methods. This only shows that these two groups of respondents agree that teachers are competent as regards these variables.

➤ *Test of Significant Difference Between and Among the Science Teachers' Test Construction Competencies when they are Grouped According to their Demographic Profile*

Table 15 presents the results of the correlation analyses which were performed to determine the relationship between and among the science teachers' test construction competencies when they are grouped according to their demographic profile.

Specified from the table that a significant difference was found between and among the test construction competencies of the Science teachers in terms of administering, scoring, and interpreting results when they are grouped according to sex ( $p=0.044$ ), highest educational attainment ( $p=0.038$ ) and seminars attended related to test constructions ( $p=0.021$ ). This significant difference is manifested by the computed probability values for these variables which are greater than the 0.01 significance level. Results of the analyses also imply that female Science teachers had a greater level of test construction competencies in terms of administering, scoring, and interpreting results. Further, Science teachers who took advanced studies such as master's and doctorate degrees had greater competencies in the area of test constructions. Additionally, Science teachers who attended seminars in test construction had a greater level of competencies in administering, scoring, and interpreting results.

Table 15 Results of the F/t-test Analysis on the Difference between and among the Science Teachers' Test Construction Competencies when they are Grouped According to their Demographic Profile

Test Construction Competencies	Demographic Profile				
	Age	Sex	Educ. Attain	Specia-Lization	Seminars
Developing Assessment Method Appropriate For Instructional Decisions	0.322 (0.214)	0.874 (0.142)	0.751 (0.237)	0.457 (0.389)	0.521 (0.421)
Administering, Scoring And Interpreting Results	0.566 (0.458)	2.237* (0.044)	2.358* (0.038)	0.321 (0.740)	2.448* (0.021)
Using Assessment Results	0.358 (0.624)	0.417 (0.551)	0.787 (0.124)	0.657 (0.221)	0.525 (0.341)
Developing Valid Students Grading Procedures	0.452 (0.551)	0.327 (0.618)	0.521 (0.428)	0.312 (0.841)	0.619 (0.328)
Communicating Assessment Results	0.249 (0.322)	0.511 (0.482)	0.741 (0.217)	0.428 (0.514)	0.387 (0.401)
Recognizing Unethical And Illegal Assessment Methods	0.361 (0.623)	0.819 (0.127)	0.781 (0.211)	0.707 (0.312)	0.678 (0.412)

Legend: \*\* = highly significant ( $p \leq 0.01$ ) Numbers in the upper entry are t/F-test values ns = Not Significant ( $p > 0.05$ ) Numbers enclosed in parentheses are probability values (p-values)

The results imply that professional development is very important for teachers to be more knowledgeable in test constructions. Further examination of the table reveals that the other demographic profile such as age and major of specialization has nothing to do with the Science teachers' test construction competencies.

➤ *Test of Significant Relationship between the Teachers' Test Construction Competencies and their Students' Academic Performance in Science*

Table 16 displays the results of the correlation analysis which was done to determine if a significant relationship existed between the teachers' test construction competencies and their students' academic performance in science.

Table 16 Results of Correlation Analysis on the Relationship between Teachers' Test Construction Competencies and their Academic Performance in Science

Test Construction Competencies	Students' Academic Performance In Science	
	R-Value	P-Value
Developing Assessment Method Appropriate For Instructional Decisions	0.632**	0.000
Administering, Scoring And Interpreting Results	0.799**	0.000
Using Assessment Results	0.576**	0.000
Developing Valid Students Grading Procedures	0.761**	0.000
Communicating Assessment Results	0.741**	0.000
Recognizing Unethical And Illegal Assessment Methods	0.634**	0.000

Legend: \*\* = highly significant ( $p \leq 0.01$ )

Demonstrated from the table that a highly significant relationship was found between the teachers' test construction competencies and their students' academic performance in Science. This highly significant relationship was brought about by the fact that the computed probability value of 0.000 for these variables is lower than the 0.01 significance level. Further observation of the table shows that a direct relationship existed between these variables as indicated by the positive sign of the computed correlation values that ranged from 0.576 to 0.799. This means that as the level of the teachers' test construction competencies increases, their students' academic performance in Science also increases.

The results imply that when teachers have adequate knowledge in designing and utilizing assessment tests, their students will be able to attain higher grades in Science.

In conjunction with the findings of the present study, Quansah and Amoako (2018) found that Senior High School teachers have a negative attitude toward test construction. The authors specifically found a poor attitude of teachers in the planning of tests, item writing, item review, and assembling of the items. Quansah and Amoako concluded that this attitude of teachers influenced the quality of tests used for assessing students. It is of essence to state that the poor attitude might not be due to their inadequate skills but also to the fact that some teachers see test construction as a burden. If objective and accurate data are to be acquired from students during the teaching and learning process, analyzing teachers' test-construction abilities is important. Moreover, the results of their study showed that students obtained lower grades which were due to poor test construction by the teacher.

During the interview that was carried out, the teachers were asked whether their experience in assessing students would affect their academic performance. The teachers believe that their level of subject-matter expertise has a direct impact on the academic progress of their students. Others have said that they go above and beyond what is expected of them as teachers because they know how to evaluate students well.

On the other hand, in the conducted interview with the students, they were asked, "How competent is your teacher in formulating questions in examinations?" the students responded that their teachers are very competent in formulating questions. Some added that their teacher is very happy when the students understand the questions and answer them correctly. A few said that their teacher sometimes gives questions that are already used in the modules.

Lastly, both teachers and students were asked, "How important are the teachers' test construction competencies on students' academic performance in Science?" Many teachers responded that it is very important for them to be competent as regards test construction because giving high-quality tests to the students motivates them to do better in exams thus affecting their grades positively. Other teachers said that teachers' test construction competencies can make or break the students' confidence in answering exams. Some added that correctly assessing students' work and outputs will keep them motivated and active in class.

In the meantime, answering the same question asked with the teachers, "How important are the teachers' test construction competencies on students' academic performance in Science?" students responded that it is very important for them that their teachers possess high competencies in test construction. According to them, it is crucial because if the tests are hard to understand and the instructions were not clear, it will be very difficult for them to get the correct answers.



➤ *Program of Activities Crafted from the Results of the Study*

Results of the study revealed that test construction in terms of recognizing unethical and illegal assessment methods yielded the lowest assessments from the teachers and their respective school heads. Further findings of the study revealed that there are some Science teachers who did not attend any seminars related to test construction. Hence, the researcher offers the Program of Activities which is presented in Table 17.

Table 17 Proposed Program of Activities

<b>Objectives</b>	<b>Action</b>	<b>Timeline</b>	<b>Persons Involved</b>	<b>Expected Outcome</b>
To dedicate one learning action cell (LAC) session to discussing exam malpractices such as cheating, leakage, falsification of exam results, etc.	Learning Action Cell (LAC) session dedicated to test construction, among science teachers	4 <sup>th</sup> Quarter of S.Y. 2022-2023	School Head, Researcher, Head Teacher Teachers	At the end of the LAC session, teachers are expected to determine different exam malpractices.
To provide lectures to science teachers focusing on the latest trends in test construction.	Invite Science experts for the In-Service training, who can share current trends and pedagogy in test construction	July 2023 (during the In-service training)	School Head, Researcher, Head Teacher, Teachers, Invited speakers	Teachers are expected to enjoy and appreciate the benefits of developing their test construction competencies.
To establish a Test Item Bank consisting of valid test items.	Establishment of Test Item Bank	July-August 2023 (before the opening of S.Y. 2023-2024)	Researcher, Head Teacher, Teachers	Teachers are expected to create valid and measurable test items that would result in a meaningful assessment for the students.

## CHAPTER FOUR FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

This chapter presents the summary of the major findings, the conclusions arrived at based on the findings, and the recommendations given in accordance with the conclusions.

### ➤ *Findings*

This study determined the influence of Grades 7 and 8 public high school teachers' test construction competencies on the students' academic performance in Science in public high schools in San Miguel, Bulacan during the School Year 2022-2023.

Using the procedures described in the preceding chapter, the answers to the problems raised in this study were ascertained and summarized as follows: Findings revealed that majority of the Science teachers belong to the age group of 24 to 38 years old, master's degree holders, female, Biology major, and attended seminars related to test construction.

The school heads and the Science teachers themselves "almost always" to "always" believed that teachers are competent in test construction in terms of developing assessment method appropriate for instructional decisions; administering, scoring, and interpreting results; using assessment results; developing valid students grading procedures; communicating assessment results; and recognizing unethical and illegal assessment methods.

The academic performance of the junior high school students in Science was described as "satisfactory".

Highly significant difference was found between the assessments of the school heads for the teachers and the Science teachers themselves with regard to their test construction competencies in terms of developing assessment method appropriate for instructional decisions. Science teacher respondents had greater assessments as compared to their school heads in so far as test construction competencies in terms of developing assessment method appropriate for instructional decisions.

Significant difference was found between and among the test construction competencies of the Science teachers in terms of administering, scoring and interpreting results when they are grouped according to sex, highest educational attainment and seminars attended related to test constructions.

Science teachers who are female, took advanced studies such as master's and doctorate degrees, and attended seminars in test constructions had greater level of test construction competencies in terms of administering, scoring, and interpreting results.

A highly significant relationship was found between the teachers' test construction competencies and their students' academic performance in science.

During the conducted interview, both teachers and students were asked, "How important are the teachers' test construction competencies on students' academic performance in Science?" These two groups of respondents claimed that the teachers' ability to construct tests greatly affects their students grades in science. Many teachers said that their ability to design and use assessment exams can affect their students' confidence in answering the test. They stated that if the students cannot comprehend or if the directions are vague, then the students might not be able to answer correctly and will lose concentration. Meanwhile, the students said that it helped them greatly when tests directions are easy to understand. Some of the students also stated that multiple choice questions created by their teachers are the easiest to answer because they have choices to choose from which help them recall past lessons.

### ➤ *Conclusions*

Based on the findings of the study, the following conclusions were drawn: Science teachers had greater assessments of themselves as compared to the assessment of school heads in test construction competencies in terms of developing assessment method appropriate for instructional decisions. Science teachers who are female, finished master's and doctorate degrees, and attended seminars in test constructions had greater level of test construction competencies in terms of administering, scoring, and interpreting results. Teachers' test construction competencies are positively correlated to students' academic performance in science.

### ➤ *Recommendations*

Considering the findings and conclusions of the study, the following recommendations are hereby offered:

- School officials may motivate and encourage the Science teachers under their supervision to pursue and enroll in graduate programs by the next school year.
- School officials may include in their In-service training and Learning Action Cell (LAC) sessions the basic skills as well as the latest trends and pedagogy in test construction,

- The program of activities offered by the researcher may be considered by the school officials for enriching the teachers and improving students' performance in Science.
- For future researchers, further research along this line could be conducted. Same study could be conducted in the senior high school to further validate the significance of the variables under study to the academic outcomes of the students in Science.

## REFERENCES

- [1]. Ahmed, K.A.E. (2018). Training In-Service Teachers in Test Construction Skills (TCS). *King Khalid University Journals*, 202-49-69
- [2]. Agu, N., Onyekuba, C., & Anyichie, A. (2013). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill inventory. *Educational Research and Reviews*, 8(8), 431-439
- [3]. Amani, J., Kitta, S., Kapinga, O., & Mbilinyi, C. (2021). Secondary School Teacher's Knowledge on Procedures for Constructing Quality Classroom Tests in Tanzania. *Universitepark Bülten*, 10(1): 40-54.
- [4]. American Federation of Teachers. (1990). Standards for Teacher Competence in Educational Assessment of Students. *Educational Measurement: Issues and Practice*, 9(4), 30–30. doi:10.1111/j.1745-3992.1990.tb00391.x
- [5]. Berame JS, Lawsin NLP, Gacosta H, Yadao M, Orleans AV. Development and Validation of Organismal Classification Achievement Test (OCAT): Basis for Implementation and Standardization for Test Construction in Junior Science High School. *Int J Biotechnol Recent Adv*. 2018; 1(1): 30-34.doi: 10.18689/ijbr-1000106
- [6]. Berman E. (2017). An Exploratory Sequential Mixed Methods Approach to Understanding Researchers' Data Management Practices at UVM: Integrated Findings to Develop Research Data Services. *Journal of eScience Librarianship*, 6(1) <https://doi.org/10.7191/jeslib.2017.1104>.
- [7]. Bernard, H. R. (2002). Research methods in anthropology: Qualitative and quantitative approaches (3rd ed.). *Walnut Creek, CA: Alta Mira Press*.
- [8]. Bongco, R. T. & Ancho, I. V. (2020). HisStory in the Feminized Teaching Profession in the Philippines. *Journal of Contemporary Eastern Asia Vol. 19, No. 2: 197-215*
- [9]. Brabeck, M. M., Dwyer, C. A., Geisinger, K. F., Marx, R. W., Noell, G. H., Pianta, R. C., ... & Worrell, F. C. (2016). Assessing the assessments of teacher preparation. *Theory into practice*, 55(2), 160-167.
- [10]. Brown, G. (2019). Is Assessment for Learning Really Assessment? *Frontiers in Education*, 4, 64.
- [11]. Clay, B. (2001). Is this a trick question? A short guide to writing effective test questions. *Kansas Curriculum Center*.
- [12]. Cristobal, J. (2022). Test Construction Principles Violated in Selected Summative Tests in English. [*Mariano Marcos State University*]. *American Journal of Arts and Sciences*, 1(2), 33-39
- [13]. Cook, D. A., Brydges, R., Ginsburg, S., & Hatala, R. (2015). A contemporary approach to validity arguments: a practical guide to Kane's framework. *Medical education*, 49(6), 560-575.
- [14]. Data.worldbank.org. (n.d.). Secondary education, teachers (% female) - Philippines, India. Retrieved from <https://data.worldbank.org/indicator/SE.SEC.TCHR.FE.ZS?locations=PH-IN>. Retrieved December 14, 2022
- [15]. David, C.C., Albert J.R.G., & Vizmanos, F.V. (2019). Toward an Adaptive Social Protection in the Philippines. *Philippine Institute of Development Studies*. ISSN 2508-0865 (electronic) No. 2019-01
- [16]. Department of Education (2015). Policy Guidelines on Classroom Assessment for the K to 12 Basic Education Program (DepEd Order No.8, s.2015). *Manila: Department of Education, Republic of the Philippines*.
- [17]. Department of Education (2016). The Learning Action Cell as a K to 12 Basic Education Program School-Based Continuing Professional Development Strategy for the Improvement of Teaching and Learning (DepEd Order No.35, s.2016). *Manila: Department of Education, Republic of the Philippines*.
- [18]. Department of Education (2020). Policy Guidelines for the Provision of Learning Resources in the Implementation of the Basic Education Learning Continuity Plan (DepEd Order No.18, s.2020). *Manila: Department of Education, Republic of the Philippines*.
- [19]. Department of Education (2020). Interim Guidelines for Assessment and Grading in Light of the Basic Education Learning Continuity Plan (DepEd Order No.31, s.2020). *Manila: Department of Education, Republic of the Philippines*.
- [20]. Dewey, R. (2003). Writing Multiple Choice Items Which Require Comprehension. *Georgia Southern University, Statesboro, GA*.
- [21]. Duvie, A. N., & Eluwa, B. O. (2016). Assessment of the effectiveness of management strategies for curbing examination malpractices in secondary schools in Nigeria. *European Journal of Education Studies*.
- [22]. Earl, L. (2003.) Assessment as Learning: Using Classroom Assessment to Maximize Student Learning. *Corwin Press*.
- [23]. Esguerra, D.J. (2018). DepED urged to lighten teacher workloads following suicide reports. *Philippine Daily Inquirer*. August 27, 2018
- [24]. Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5 (1):1 DOI:10.11648/j.ajtas.20160501.11
- [25]. Fauth, B., Decristan, J., Decker, A. T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education*, 86, 102882.
- [26]. Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, 115(4), 523-547.
- [27]. Hakim, A. (2015). Contribution of competence teacher (pedagogical, personality, professional competence and social) on the performance of learning. *The International Journal of Engineering and Science*, 4(2), 1-12.
- [28]. Haladyna, T. (2018). Developing Test Items for Course Examination. *IDEA Center Inc*.
- [29]. Hardcastle J., Herrmann-Abell C.F., DeBoer G.E. (2017). Comparing Student Performance on Paper-and-Pencil and Computer-Based-Tests. *American Association for the Advancement of Science*.

- [30]. Hamafyelto, R.S., Hamman-Tukur A., & Hamafyelto S.S. (2015). Assessing Teacher Competence in Test Construction and Content Validity of Teacher Made Examination Questions in Commerce in Borno State, Nigeria. *Scientific & Academic Publishing*, 5(5), 123-128
- [31]. Hobbs, L. (2015). Too many teachers teaching outside their area of expertise. Retrieved from <https://theconversation.com/too-many-teachers-teaching-outside-their-area-of-expertise-39688>. Retrieved December 14, 2022
- [32]. Ibad, F. (2018). Personality and Ability Traits of Teachers: Student Perceptions. *Journal of Education and Educational Development*, 5 (2), 162 – 177
- [33]. Jones, C. (2005). Assessment for learning. *Learning and Skills Development Agency*.
- [34]. Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169-197.
- [35]. Magno, C. (2013). Standards of Teacher Competence on Student Assessment in the Philippines. *The Assessment Handbook*, 10, 42-53
- [36]. Miracle, S. (2021). The Shifting Role of Teachers: A Case Study on Teachers' Perceptions of Personalized Learning in a Rural School District. *Doctoral Dissertation, University of Louisville*.
- [37]. Oregon State University. N.D. Classroom Assessment Tools. *Center for Teaching and Learning*.
- [38]. Quansah, F., & Amoako, I. (2018). Attitude of Senior High School (SHS) teachers towards test construction: Developing and validating a standardized instrument. *Research on Humanities and Social Sciences*, 8(1), 25-30.
- [39]. Rahman, S., & Majumder, M. (2015). Is it Assessment of Learning or Assessment for Learning? *South East Asia Journal of Public Health*, 4 (1).
- [40]. Rasyidah, U., Triana, N., & Saukah, A. (2020). The Teachers' Assessment Knowledge and Practices: Contributions of the Past-Time Experiences to the Present-Time Decision. *The Qualitative Report*, 25(7).
- [41]. Rawluysk, P.E. (2018). Assessment in Higher Education and Student Learning. *Journal of Instructional Pedagogies*, 21.
- [42]. Roelofs, Erik & Cito,. (2022). Imaging and assessing teacher competence: design characteristics of structured video portfolios.
- [43]. Shava, G., Hleza, S., Tlou, F., SHonhiwa, & Mathonsi, E. (2021). Qualitative Content Analysis, Utility, Usability and Processes in Educational Research. *International Journal of Research and Innovation in Social Science*, 5 (7).
- [44]. Shallianah, S. (2016). An Analysis of the Teacher Competence on Developing Rubric to Assess English Competence of Eight-Grade Students in SMPN 1 Sedati Sidoarjo. *Thesis. Sunan Ampel State Islamic University*.
- [45]. Sohrab S., Mohammad S.B., Bahram E. (2012). Studying the Relationship Between Students' Performance in Scientific Examination and Their Academic Achievement in Third Year of High School. *International Journal of Science and Engineering Investigations*, 1(7)
- [46]. Taherdoost, H. (2016). Sampling Methods in Research Methodology; How to choose a sampling technique for research. *International Journal of Academic Research in Management*, 5(2), 18-27
- [47]. Tran N. (2014). The Impact of Assessment on the Learners' Identities: A Literature Review. *ARECLS Vol.11, 2014, 90 - 106. 90* <https://research.ncl.ac.uk/media/sites/researchwebsites/arecls/The%20Impact%20of%20Assessment%20on%20the%20Learners%20Identities%20A%20Literature%20Review.pdf>
- [48]. Tschannen-Moran, M., & Gareis, C. R. (2015). Faculty trust in the principal: An essential ingredient in high-performing schools. *Journal of Educational Administration*, 53(1), 66-92.
- [49]. Ughamadu, U., & Ifeyinwa, A.P., & Okaforcha, C. (2021). Validation of Test Construction Skill Inventory for Quality Assurance in Colleges of Education: Implications for Educational Management. *COOU Journal of Educational Research*, 6(1), 77-91.
- [50]. Veldkamp, B.P. (2015). Test Construction. *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*.