Analysis on Teachers' Summative Assessment in Basic Secondary Education

Raul C. Orongan[1*], Alven A. Manual[2], Maris Jade Q Orongan[3], Lorelie V Gamutan[4], Jayvy C Vegafria[5]

[1]Professional Education Department, College of Education, Central Mindanao University, Musuan, Bukidnon, Philippines 8714
[2]Secondary Education Department, College of Education, Central Mindanao University, Musuan, Bukidnon, Philippines 8714
[3]Science Education Department, College of Education, Central Mindanao University, Musuan, Bukidnon, Philippines 8714
[4,5]Department of Education, Valencia City Division, Bukidnon, Philippines
[1*] f.raul.orongan@cmu.edu.ph,

Abstract. This study assessed the different summative tests of teachers given to learners in Grade 11 students' English, Filipino, Mathematics, and Science subjects as listed in the National Achievement Test. Specifically, it aimed to identify the learning competencies in the Table of Specifications (TOS) with the behavioral skills in the test in the four subject areas, determine the reliability of the summative tests, and ascertain the match between the learning competencies with the retained tests per subject area. The research utilized the ex post facto design. Secondary data were obtained from the TOS and summative tests. Results showed that the TOS specified learning objectives using a division-wide format. The desired skills were classified as easy, moderate, and challenging levels. The Cronbach Alpha values in the summative tests of the four subject areas were found to be of acceptable reliability. The item analysis summary results indicated the percentage of test items retained in the specified learning competency with the prescribed skill level in all four subject areas. A reliable test was indicated by the proportion of the test's total items to its computed retained items, as determined by the TOS's cognitive behavior. The things that were kept included some that were of a lesser skill level.

Keywords: Summative Assessment, Learning Competencies, Table of Specifications, Reliability, Item Analysis

1 Introduction

Throughout the academic year, educators at all levels of education prepare and deliver a large number of official teacher-made examinations. For this reason, tests are

J. Handhika et al. (eds.), *Proceedings of the 4th International Conference on Education and Technology (ICETECH 2023)*, Atlantis Highlights in Social Sciences, Education and Humanities 25,
https://doi.org/10.2991/978-94-6463-554-6_36

essential tools in the educational process. Adherence to the principles of test development, administration, analysis, and reporting is crucial, particularly in the context of developing norm-referenced assessments for educational objectives. The test is the most essential part of the learning process. Most teachers are not concerned with the quality standard when designing tests. They only fulfill their obligations to design, apply, and score the test. The government intended for the Enhanced Basic Education Law, or Republic Act 10533, to create a graduate workforce capable of meeting international standards.

However, in gauging quality education, schools would adhere to the minimum standards set by the Department of Education, which is the National Achievement Test (NAT) administered to learners in Grades 6, 10, and 12. Valencia City Division, in particular, is not exempted from the minimum results of their learners. The results showed that the low performance of NAT learners in Valencia City calls for immediate attention from administrators and stakeholders. The shift to the K to 12 curriculum promised to allow every student to receive a globally competitive quality education based on a pedagogically sound curriculum at par with international standards (RA 10933).

The College of Education, Central Mindanao University, as one leading institution in the region and as a Center of Development, responds to this call for complementing primary and tertiary education. Currently, the college is conducting an extension program for the teachers of the Valencia City Division. One of the components of the extension is assessing the teaching and learning process by which this study is anchored. The Division is adopting one type of summative assessment at all levels using the Multiple-Choice Test (MCT) in the summative assessment in the quarter examinations. Accordingly, Though they have more ability to write than memory-based items, well-written MC items can evaluate higher-level cognitive processes. Also, more effort and time are required to construct good-quality MCTs than other questions. This is only possible if the test constructor follows rigidly the numerous guidelines for writing MCTs.

Thus, this study simultaneously assessed the different summative tests made by the basic secondary education teachers in their quarterly examinations, which would match the minimum learning competencies of learners to what has been taught and be assessed based on specified standards. The results of the study are beneficial to the target clientele, specifically: a) teachers – as the facilitator of the class in balancing the teaching-learning process as manifested by the results of the summative tests; b) DepEd administrators/curriculum implementers – in evaluating the quality of instructional interface using assessment results for policy review and revision; c) students – in achieving the maximum learning competencies which are at par with the standards set at their norm level; and d) College of Education Administrators/Teachers – for partnership with the Department of Education in boosting quality instruction in secondary level thru Research, Development and Extension (RDE) collaboration.

2 Review of Literature

An essential component of implementing the program is classroom assessment. It enables educators to monitor and assess students' development and modify their lessons accordingly. Students' academic achievement in school is communicated to parents and guardians through classroom assessments. So, in terms of overseeing pupils' learning, teachers were at the forefront. Since the new curriculum was introduced in grades K–12, it is crucial, especially in the basic education sector, to support and evaluate the teaching-learning process by the year-specific curriculum guides. One of the most important aspects of education is assessment, which measures the extent to which predetermined objectives have been met and provides teachers with feedback on how to improve their methods of instruction. One of the key elements of excellent teaching and learning settings is assessment.

The Philippine educational system has come to light as a result of the latest exams given to students both locally and internationally. According to the 2019 National Achievement Test (NAT), students in the Philippines do poorly in the subjects of science, math, and English. This outcome is made worse by the most recent results of the Program for International Student Assessment (PISA), which showed that among the 79 participating counties, Filipino pupils came in last. The data strongly suggest a comprehensive performance assessment down to the grassroots level in secondary education. In specifics, the topics of Science, Mathematics, English, Filipino, and Araling Pan-lipunan all contain five (5) areas and competencies that are tested in the NAT. These five disciplines were designed to help secondary students develop their lifetime learning skills under the Restructured Basic Education Curriculum. The Secretary of the Department of Education also underlined that learning outcomes are now expected to be more than just reading, writing, and counting. The 21st-century skills covered in K–12 curricula go beyond lifelong learning abilities and encompass more than just subject matter and concept mastery. It has to do with acquiring digital literacy as well as problem-solving and critical thinking abilities. These skills are assessed in the National Achievement Test, providing observational information on the academic achievement level of secondary learners and their strengths and weaknesses in major subjects.

Assessment is connected to the learning goals of diagnosis, prediction, placement, evaluation, selection, grading, guiding, and administration. In terms of both design and implementation, student learning is the top priority when it comes to assessment for learning. As such, it is not the same as an assessment that is primarily used for ranking, accountability, or competency certification. The majority of assessment activities facilitate learning when they offer data that teachers and students may use as feedback to evaluate one another and themselves and adjust the teaching and learning activities they are participating in. This ultimately justifies the establishment of learning objectives as a basis for creating the proper instruments and techniques for evaluation and, during the learning process, determining what is worthy and necessitates Understanding. ,

considering students understanding of goals had critical motivational and cognitive impacts.

The difficulties faced by students and the secondary education system as a whole were recognized by DepEd given this fact. There is a bright side to this situation, according to DepEd, since it will act as a wake-up call for all parties involved in education to collaborate to achieve better education. Therefore, it is recommended that State Universities and Colleges (SUCs) take part in the evaluation of basic education students to determine their needs and competencies comprehensively. Since secondary students attending higher education institutions (HEIs) are expected to demonstrate fundamental educational competencies, the evaluation is based on this reality. Based on the reports, In the Philippines, there have been notable advancements in teacher preparation, classroom assessment, and basic education over the last five years. In light of these reforms, the teacher education curriculum developed by the Commission on Higher Education (CHED) must be evaluated to make sure it aligns with the new initiatives.

The HEIs are called to assist in addressing the gaps in the secondary educational system by comprehensively evaluating the student's academic performance. The academic assessment is timely and essential, as this will focus on the experiences and perceptions of secondary students. This endeavor envisions assisting DepED in enhancing students' performance in national achievement tests. This research is also based on Bloom's taxonomy's cognitive level. that knowledge is acquired in a methodical, step-by-step process that begins with lower-order cognitive abilities (LOTS) like recall and comprehension and progresses to higher-order cognitive abilities (HOTS) like assessing, generating, and analyzing

3 Objectives

This study aimed to assess the different summative tests teachers give to learners in the subject areas specified in the National Achievement Test. Specifically:

1. Identify the learning competencies in the summative test made by teachers based on the Table of specifications (TOS) in the following subject areas:
    a. English
    b. Filipino
    c. Mathematics
    d. Science
2. assess the reliability indices of the given summative tests in the specified subject areas and
3. analyze the result of test items concerning the learning competencies/ objectives per subject area.

4 Methodology

This is a quantitative study utilizing ex post facto design. Secondary data were obtained from the summative scores of Grade 11 students during the fourth grading period, which was the start of the pandemic outbreak. The study's conceptualization was the offshoot of the extension program of the College of Education and the Department of Education, Valencia City Division, under the Teaching and Learning Assessment project. Multiple types of tests were employed division-wide for the summative exam. Only four subject areas (4), namely English, Filipino Mathematics, and Science were analyzed in this study, considering these are the significant areas in the National Achievement Test (NAT).

The division coordinators requested secondary data, such as the Table of specifications per subject area. The coordinators also made available the results of the fourth grading summative test. The data were described and examined using the descriptive method. We used descriptive statistics, including percentages and frequency counts. An investigation of reliability was conducted utilizing the Cronbach Alpha index. Item analysis was also employed to evaluate the TOS's learning competencies in relation to the items that were kept on the exam after standardization.

5 Results and Discussion

Teachers can better integrate objectives, instruction, and assessment with the use of a Table of Specifications (TOS), often known as a test blueprint. This strategy can be used for various assessment methods but is most commonly associated with constructing traditional summative tests. Table 1 shows the TOS for the English subject of Grade 11 students during the fourth quarter summative examination.

The format was outlined and adopted for the whole Division of Valencia City. The first column displayed the learning objectives/ skills /topics covered for a particular grading period. For English, a total of 13 learning competencies were listed. Also found in the Table were the number of hours the topic allotment, the proportion of the test coverage, and the total number of items obtained per topic. (However, no data was reflected here in this TOS and would be supplied by the coordinator). The skills or the behavioral measures were categorized into three parts: Easy, Moderate, and Difficult, with 60, 30, and 10 percent allocation, respectively.

Table 1. TOS for English

TOPIC/S SKILLS/ OBJECTIVES	S K I L L S						Total
	60% Easy		30% Moderate		10% Difficult		
	R	U	AP	AN	E	C	

Explains the nature and process of communication	4	2			6
Distinguishes the unique feature(s) of one communication process from the other	2	3			5
Demonstrates sensitivity to the sociocultural dimension of communication situations with a focus on culture, gender, age, social status, and religion	2	2			4
Identifies the speaker's purpose	3	1			4
Identifies strategies used by each speaker to convey his/her ideas effectively	6				6
Explains that a shift in speech context, speech style, speech act, and communicative strategy affects various factors	1	5			6
Compare and contrast patterns of written texts across disciplines.		5		7	12
Explain critical reading as reasoning	1	2			3
Analyse a written text according to its mechanics, organisation, coherence, and cohesiveness.	1	1		1	3
Determine which statements in a textual text are made directly or indirectly.		1		1	2
Create evaluations based on the text you've read.	1	2			3
Determine the special characteristics and needs for writing texts that are applicable to several academic fields.	2	3			5
Determine the setting in which a text was created.	1				1
TOTAL	**24**	**27**	**0**	**9**	**60**

(R-Remembering; U-Understanding; AP-Applying; AN-Analyzing; E-Evaluating; C-Creating)

This implies that when designing the Table of specifications, it is vital to consider that the outline must relate to the target behavior as instruction is implemented based on the set learning standards. Through it, teachers can determine what topic should be stressed out. This also assists in preparing tests that reflect what students have learned and limits the amount of time spent on each unit, albeit with different approaches in developing and using a TOS advocated by measurement experts.

Table 2 depicts the TOS for Filipinos, with a format similar to Table 1, showing the number of hours allotted for the particular learning competencies. The Table serves to define the scope and focus of the test clearly. It ensures that the teachers include items with different levels of cognitive complexity when measuring students' achievement so that they will remember the details.

Table 2. TOS for Filipino (*National Language*)

TOPIC/S SKILLS/ OBJECTIVES *(PAKSA/KASNAYANG PAMPAGKATUTO)*	SKILLS / KASANAYAN						Total
	60% Easy (*Madali)*		30% Moderate / *(Katamtaman)*		10% Difficult (*Mahirap)*		
	R	U	Ap	An	E	C	
Nasusuri ang ilang halimbawang pananalikslk sa FIiipino batay sa layunin, gamit, metodo, at etika sa pananaliksik	2	1	1	1	0		5
Nabibigyang kahulugan ang mga konseptong kaugnay pananalikslk, (halimbawa: balangkas kenseptuwal, balangkas, teoritikal, datos empirical, atbp	3	10	1	3	2		19
Nailsa-isa ang mga paraan at tamang proseso ng pagsuJat ng isang pananaliksik sa Filipino batay sa layunin, gamit, metodo, at etika ng pananliksik	2	11	4	3	1		21
Nagagamit ang mga katuwirang lohikal at ugnayan ng mga ideya sa pagsulat ng isang pananaliksik	0	4	1	1	0		6
Nakabubuo ng isang maikling pananaliksik na napapanahon ang paksa.	1	2	2	1	0	3	9
TOTAL	**8**	**28**	**9**	**9**	**3**	**3**	**60**

(R-Remembering; U-Understanding; AP-Applying; AN-Analyzing; E-Evaluating; C-Creating)

Table 3 shows about twenty-six (26) learning objectives/competencies for Mathematics TOS. It also notes that no items were indicated in the difficult level. This only suggests that students' thinking ability in solving problems in mathematics will be reflected in their ability to solve the given problem. Therefore, the steps to solve the mathematical test tend to be unlimited, and the differences in sequences depend on the student's ability to master mathematical concepts.

Table 3. TOS for Mathematics

TOPIC/S SKILLS/ OBJECTIVES	SKILLS						Total
	60% Easy		30% Moderate		10% Difficult		
	R	U	AP	AN	E	C	
FUNCTION Represents real-life situations using piece-wise functions, Operation of Functions, and solving problems involving functions.	2						2

Solves rational equations and inequalities; distinguishes rational functions, rational equations, and rational inequality;	1	1	1			3
Computes interest, compound interest and simple interest, maturity value, future value, and present value.		1		1		2
Demonstrates an exponential function with its a) equation, b) graph, and c) table of values	1					1
Resolves exponential inequality and equation problems.	1	1				2
Uses one-to-one functions to simulate real-world scenarios.	2	1				3
ascertains a one-to-one function's inverse	1	1	1			3
Differentiates exponential inequality, exponential equation, and exponential function	1					1
incorporates logarithmic functions to represent real-world scenarios.	2		1			3
Differentiates between interests that are simple and complex	1					1
Distinguishes between a logarithmic function, equation, and inequality	1			1		2
demonstrates the logarithmic laws				1		1
Uses exponential functions to represent scenarios found in real life.	1	2				3
Differentiates exponential inequality, exponential equations, and exponential functions	2					2
Depicts a random variable (both continuous and discrete).	2					2
Illustrates a probability distribution for a discrete random variable and its properties	1		1			2
Interprets a discrete random variable's mean and variance,	1		1			2
Solves probability distribution problems requiring mean and variance	2	1				3
Distinguishes between random variables that are discrete and continuous.	2	1				3
Determines the statistical sampling distributions (sample mean).	1	1				2
Determines a confidence interval's length.	1	2				3
Determines the right sample size by calculating the interval's length.	1	1				2
Identifies the parameter to be tested given a real-life problem.	3			1		4
Based on the test-statistic result and the rejection region, get the population mean.	1					1
resolves issues with sample mean sampling distributions.	1	2		2		4
calculates the confidence interval's length.		2	1			3
Total	**32**	**17**	**6**	**5**	**0**	**60**

(R-Remembering; U-Understanding; AP-Applying; AN-Analyzing; E-Evaluating; C-Creating)

Table 4 reflects the science subject area for Grade 11; the topics were subdivided into two divisions, Earth Science and Physical Science, with 17 and 19 learning competencies, respectively, and 36 objectives. The science teachers prepared 60 test items. Each subject composition contains 30 items for the summative test. The distribution of the test items was so diverse that it mainly contained one question per learning skill.

Table 4. TOS for Science

TOPIC/S SKILLS/ OBJECTIVES	SKILLS						Total
	60% Easy		30% Moderate		10% Difficult		
	R	U	AP	AN	E	C	
EARTH SCIENCE							

TOPIC/S SKILLS/ OBJECTIVES	SKILLS						Total
	60% Easy		30% Moderate		10% Difficult		
	R	U	AP	AN	E	C	
Acknowledge that Earth is special because it is the only planet in the solar system with the conditions needed to support life.	1						1
Describe the four subsystems that make up the Earth and the boundaries that matter and energy flow through.		1					1
Sort rocks into three categories: metamorphic, sedimentary, and igneous.		2					2
elucidate how erosion carries weathering products away and deposits them somewhere else.	1		1				2
Examine and contrast how the various igneous rock types were formed.			2				1
Describe how folds and flaws are created by the movement of plates.	1		1				2
Explain the various techniques (absolute and relative dating) used to establish a stratified rock's age.	1						1
Describe the process by which the divisions of geologic time were determined using relative and absolute dating.			1				1
Explain how the geologic time scale can be used to understand the history of the Earth.	1						1
Explain the different risks that might arise from landslides, volcanic eruptions, and earthquakes.		3					3
Explain how the idea of life is changing in light of new information.	1		1				2
Describe the ways in which the study of life has unified themes (such as structure and function and ecosystems).	1						1
Consider the advantages and drawbacks of utilising GMOs.	2				1		3
Examine and understand how the various organ systems interact to ensure the survival of the animal.				3			3
Describe the ways in which organism populations have changed and will continue to change over time.		2					2
Describe the ways in which evolutionary relationships form the basis of the current classification system for organisms.	2						2
SUB TOTAL (EARTH SCIENCE)	**12**	**8**	**6**	**3**	**1**	**0**	**30**
PHYSICAL SCIENCE							
Provide proof and an explanation of how heavier elements develop during the formation and evolution of stars.	1						1
Describe how the discovery of new elements in the lab was made possible by the idea of atomic number.		2					1
Based on a molecule's structure, determine if it is polar or nonpolar.	1						1
Connect a molecule's polarity to its characteristics.			1				1
What are the common categories of intermolecular forces?	1						1
Describe how the intermolecular forces affect a substance's qualities.			1				1
Describe how the properties and functions of biological macromolecules are determined by their structures.		1					1
Explain how temperature, particle size, and concentration affect reaction rate using basic collision theory.			1				1
Calculate the amount of product created in a reaction by identifying the limiting reactant.	2						2
Explain the process of obtaining energy from various sources.		2					2
Justify the Greeks' understanding of the Earth's spherical shape.	1		1				2
Describe how Brahe's massive data collecting and innovations work.		1	1				2
Explain how Galileo inferred that an object in a vacuum falls.		1		1			2

TOPIC/S SKILLS/ OBJECTIVES	SKILLS						Total
	60% Easy		30% Moderate		10% Difficult		
	R	U	AP	AN	E	C	
Describe the applications of the photon concept and the fact that a photon's energy is directly related to its frequency.	1			1			2
Provide an experimental demonstration of electrons' wave-like behavior.			2				2
Describe how Maxwell's electromagnetic theory and Newtonian mechanics were reconciled by special relativity.		1		2			3
Describe the effects of Special Relativity's postulates.				2			2
Describe the methods used to measure the distances and speeds of distant objects, such as the cosmic distance ladder and the Doppler effect.				1			1
Describe how our universe, which is estimated to be 14 billion years old and was once hot, is expanding and how we exist in it.				1			1
SUB TOTAL (PHYSICAL)	**7**	**8**	**7**	**8**	**0**	**0**	**30**
TOTAL	**19**	**16**	**13**	**11**	**1**	**0**	**60**

(R-Remembering; U-Understanding; AP-Applying; AN-Analyzing; E-Evaluating; C-Creating)

Classroom tests provide teachers with essential information to decide instruction and student grades. A table of specification (TOS) can help teachers frame the decision-making process of test construction and improve the validity of teachers' evaluations based on tests constructed for classroom use. Mehrens and Lehmann identify that “specs” can help to provide for optimal learning on the part of students and optimal teaching efficiency on the part of the teacher. The cornerstone of classroom assessment practices is the validity of the judgments about students’ learning and knowledge. A TOS is one tool teachers can use to support their professional judgment when creating or selecting tests for their students. The TOS can be used in conjunction with lesson and unit planning to help teachers make clear the connections between planning, instruction, and assessment.

Table 5 presents the reliability indices of the four (4) subject areas, namely English, Filipino, Mathematics, and Science of Grade 11 students. This data indicates a measure to assess the reliability or internal consistency of a set of scales or test items. Cronbach's alpha is the most widely used method for estimating the internal strength of that consistency. Based on the Cronbach alpha standard values of the different summative tests, Mathematics obtained the highest alpha value of 0.79, with Filipino as the lowest 0.72 alpha value. However, all subject areas belong to the "acceptable" range of internal consistency-based reliability. This indicates that the summative division-wide test is reliable. Indicators of an acceptable, sufficient, or satisfactory level for alpha were also provided by some publications as a threshold or cut-off. These were commonly regarded as acceptable values, with values ≥0.70 or >0.70".

Table 5. Reliability indices of the summative test per subject area

Subject Area	No. of test items	Cronbach Alpha value	Qualitative Interpretation
ENGLISH	60	0.762	Acceptable Reliability
FILIPINO	60	0.723	Acceptable Reliability
MATHEMATICS	60	0.792	Acceptable Reliability
SCIENCE	60	0.764	Acceptable Reliability

(Legend: 0.90-1.00, Excellent Reliability; 0.80-0.89, Good Reliability; 0.70-0.79, Acceptable Reliability; 0.60-0.69, Questionable Reliability; 0.50-0.59, Poor Reliability; and Less than 0.59, Unacceptable Reliability)

The dependability of the test results is one of the main considerations while creating multiple-choice questions for exams. Good multiple-choice questions are typically those that have undergone a thorough item analysis procedure. The process of gathering, compiling, and utilizing data from students' answers to evaluate the test items' quality is known as item analysis. We can examine the features of a specific item by using item analysis. It can be used to determine whether test items should be improved upon or are suitable for inclusion. Thus, Table 6 showed the items retained following the item analysis procedure in the four subject areas—English, Filipino, Science, and Mathematics for Grade 11 students as well as the summary of corresponding learning competencies.

The Valencia City Division adopted three classifications with the corresponding percentage allocation in terms of skill levels categorized as Easy (60%), composed of Remembering and Understanding Cognitive domains; Moderate (30%) level with Applying and Analyzing domains; and Difficult (10%) to include Evaluating and Creating higher-order cognitive domains. As shown in Table 6, the percentage of the retained items was 56.9% and 44.4% for Easy and Moderate Multiple-Choice Questions (MCQ) type, respectively. In the English subject area, no MCQ test was given. An overall percentage of 56.7% was obtained. Take note that more items were retained at the Easy skill level.

Table 6. Summary of the items retained in the TOS learning competencies after item analysis for English, Filipino, Mathematics, and Science

SUBJECTS	SKILLS						Total No. of Items	Total No. of retained items
	Easy (Remembering/ Understanding)		Moderate (Applying/Analyzing)		Difficult (Evaluating/Creating)			
	Number of items	Retained Items	Number of items	Retained Items	Number of items	Retained Items		
ENGLISH								
Total	51	29	9	4	0	0	60	34
% Retained	*(29/51) =* ***56.9%***		*(4/9) =* ***44.4%***				*(34/60) =* ***56.7%***	
FILIPINO								
Total	36	17	18	6	6	1	60	24
% Retained	*(17/36) =* ***47.2%***		*(6/18) =* ***33.3%***		*(1/6) =* ***16.7%***		*(24/60) =* ***40.0%***	
MATHEMATICS								
Total	49	31	11	6	0	0	60	38
% Retained	*(31/49) =* ***63.3%***		*(6/11) =* ***54.5%***				*(38/60) =* ***63.3%***	
SCIENCE								
Earth Science	20	14	10	7	0	0	30	21
% Retained	*(14/20) =* ***70.0%***		*(6/9) =* ***66.7%***				*(21/30) =* ***70.0%***	
Physical Science	15	8	14	7	1	0	30	15
% Retained	*(8/15) =* ***53.3%***		*(7/14) =* ***50.0%***		*0*		*(15/30) = 50.0%*	
Total	35	22	24	14	1	0	60	36
% Retained	*(22/35) =* ***62.9%***		*(14/24) =* ***58.3%***		*0*		*(36/60) =* ***60.0%***	

In the Filipino subject area, the item analysis summary result of the summative test showed 47.2%, 33.3%, and 16.7% retained item percentages in Easy, Moderate, and Difficult skill levels, respectively. Overall, 40% of the retained items were obtained. As observed, the Filipino subject obtained the lowest overall percentage of retained items. It was further noticed that more retained items were in the Easy skill level.

The Table shows that 63.3% of the retained items were obtained in the mathematics subject area. About 63.3% and 54.5% of items were retained in the Easy and moderate levels, respectively. No MCQ test was assigned to the Difficulty level, and more test items were retained in the Easy level.

For the Science subject area, the percentage of retained items depicted an overall value of 60.0%, obtained from 62.9% and 58.3% retained items in Easy and Moderate skill levels, respectively. However, the Earth Science division has a higher percentage of item test retained, 70.0%, compared to Physical Science, with 50.0%. It was also noticed that a more significant percentage of items were retained in the Easy skill.

Using the revised Bloom's Taxonomy, lower-order thinking skills (LOTS) items retained were higher than the Higher-order thinking skills (HOTS) in all the subjects understudy. Nevertheless, 40-60% of retained tests indicate a good and reliable test. Test quality is also a crucial problem since exams are vital for providing teachers with feedback on their instructional strategies. A teacher must determine whether the test

items accurately represent the student's performance in the course concerning the particular learning objectives covered during the period after the exam has been administered and scored. Thus, educators must use evaluations to determine how well a student is meeting any given learning objective, whether it be cognitive, affective, or psychomotor. It is essential to remember that teachers' knowledge in assessment and evaluation is not a static process but rather a complex, dynamic, and ongoing activity. Classroom teachers need to update their knowledge regarding assessment practices constantly. There is a need for classroom teachers to constantly update their knowledge regarding assessment practices.

6 Conclusion

This study aimed to assess the different summative tests of teachers given to learners in the subject areas specified in the National Achievement Test, namely English, Filipino, Mathematics, and Science, for Grade 11 students. The learning objectives/competencies of the four subject areas were specified in the Table of Specifications (TOS). The Division designed a template for uniformity for the TOS and the percentage distribution of the desired skill level across the new taxonomy in the summative test. The desired skill was classified as easy, moderate, and difficult levels.

Based on the reliability analysis result, all tests in four subject areas were found to be reliable using Cronbach Alpha. Mathematics test got the highest alpha value, followed by Science, English, and Filipino. However, all four subject areas belong to the same range of reliability index classification of alpha greater than 0.70, which was interpreted as "Good for a classroom test." The item analysis summary results indicated the percentage of test items retained in the specified learning competency about the prescribed skill level in all four subject areas. A reliable test was indicated by the ratio of the test's computed retained items to its total items based on the cognitive behavior in the TOS. Additionally, a portion of the goods that were maintained at the lower skill level were.

References

1. Akem, J.A & Agbe. N.N. (2003). Rudiments of Measurement and Evaluation in Education Psychology Makurdi: The Return Press.
2. Anderson, L. W. and Krathwohl, D. R., et al. (Eds..) (2001) A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Allyn & Bacon. Boston, MA (Pearson Education Group). https://quincycollege.edu/wp-content/uploads/Anderson-and-Krathwohl_Revised-Blooms-Taxonomy.pdf
3. Black, P., C., H., C. Lee, B. Marshall & D. Wiliam (2003). Assessment for Learning – Putting it into Practice, Open University Press, Maidenhead.
4. Bloom, B. S. (1956). Taxonomy of Educational Objectives: The Classification of Education Goals. Cognitive Domain. Longman. https://sites.google.com/site/wongyauhsiung/ https://sites.google.com/site/wongyauhsiung/edu-5033-principles-and-practice-of-assessment/l2-i-learning-theories-ii-assessment-of-learning

5. Brookhart, S. M., Andolina, M., Zuza, M., & Furman, R. (2004). Minute math: An action research study of self-assessment. Educational Studies in Mathematics.
6. Buckles, S. and J.J. Siegfried, (2006). Using Multiple-Choice Questions to Evaluate In-Depth Learning of Economics, Journal of Economic Education 37(1), pp. 48-57.
7. DeMeester, K., & Jones, F.(2009). Formative assessment for P.K.–3 mathematics: A review of the literature
8. Department of Education, 2019. DepEd Chief: Teachers are still at the forefront of 21st-century learning. Retrieved at https://www.deped.gov.ph/2019/10/08/deped-chief-teachers-still-at-the-forefront-in-21st-century-learning/
9. DepEd Order #8, series 2015. Policy Guidelines on Classroom Assessment for the K to 12 Basic Education Program (BEP)
10. Derrick, J. & K. Ecclestone. (2008). "English-language Literature Review," in Teaching, Learning and Assessment for Adults: Improving Foundation Skills, OECD Publishing.
11. Fives, H & DiDonato-Barnes, N. (2013). Table of Specifications. Practical Assessment, Research & Evaluation, Vol 18, No 3 Page 2, ISSN 1531-7714
12. Griethuijsen, R. A. L. F., Eijck, M. W., Haste, H., Brok, P. J., Skinner, N. C., Mansour, N., et al. (2014). Global patterns in students' views of science and interest in science. Research in Science Education, 45(4), 581–603. doi:10.1007/s11165-014-9438-6.
13. Hopkins, C & Antes, R. (1990). Classroom Measurement and Evaluation. Itasca, Illinois, F. E. Peacock Publishers.
14. Kubiszyn. T.K. Borich.G. (2003). Educational Testing and Measurement: Classroom Application & practice.(7th ed) New York: John Wiley and Son. Inc.
15. Lamprianou, I., Athanasou, J., (2009). A Teacher's Guide to Educational Assessment. Sense Publishers: Rotterdam
16. Mehrens, W.A & Lehmann, I.J. (2009). Measurement and Evaluation in Education and Psychology. Chicago: Holt, Rinehalt, and Wonston, Inc.
17. Navarro, RL., Santos, RG.(2013). Authentic Assessment of Student Learning Outcomes Assessment of Learning 2 Second Edition. (2013). Lorimar Publishing, Inc.
18. Notar, C.E., Zuelke, D. C., Wilson, J. D. & Yunker, B. D. (2004). The Table of specifications: Insuring accountability in teacher-made tests. Journal of Instructional Psychology, 31, 115-129.
19. Noveria, A. (2018). " Item Analysis on the Validity and the Reliability of English Summative Test for the First Year Students of MA Madani Alauddin PAO-PAO, "International Journal of Management and Applied Science (IJMAS), pp. 20–23, Volume-4, Issue-5
20. Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. BMC Med Educ. 2007 Nov 28;7:49. Doi: 10.1186/1472-6920-7-49. PMID: 18045500; PMCID: PMC2148038.
21. Pandra V, Sugiman, Mardapi D. (2017). Development of Mathematics Achievement Test. International Electronic Journal Of Mathematics Education. E-ISSN: 1306-3030. 2017, VOL. 12, NO. 3, 769-776
22. Plata, S. (2018). Minding the Gap: Exploring Assessment Education in CHED's Teacher Preparation Program in the Context of PPST and DepEd's Assessment Reform. DLSU Research Congress 2018 De La Salle University, Manila, Philippines.
23. Popham, J. W. (2008). Classroom assessment: What teachers need to know (5th ed.) Boston, MA: Pearson Education, Inc. Research into Higher Education and Open University Press.
24. Quaigrain, K. & Arhin, A. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation, Cogent Education, 4:1, 1301013, DOI: 10.1080/2331186X.2017.1301013

25. Reynolds, C. R., Livingston, R. B., & Wilson, V. (2006). Measurement and Assessment in Education. Pearson: Boston.
26. Trizano-Hermosilla, I and Alvarado, J. (2016). Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. Frontiers in Psychology Journal, Vol 7, DOI=10.3389/fpsyg.2016.00769 ISSN=1664-1078
27. UNESCO. (2011). World Data on Education. 7th edition. Retrieved at https://unesdoc.unesco.org/ark:/48223/pf0000193188
28. Wiggins, G., & McTighe, J. (2000). Understanding by design. New York: Prentice Hall.
29. Wolming, S. & Wikstrom, C. (2010). The concept of validity in theory and practice. Assessment in Education: Principles, Policy & Practice, 17, 117-132.
30. Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese College EFL. TESOL Quarterly, 43, 493–513.