International Journal of Multidisciplinary Research and Growth Evaluation.

Junior High School Teachers' Conceptions and Practices of Assessment: Basis for Enhanced Faculty Development Program

John Rey O Depone
Polytechnic University of the Philippines, Philippines

* Corresponding Author: **John Rey O Depone**

Article Info
ISSN (online): 2582-7138
Volume: 06
Issue: 03
May-June 2025
Received: 07-04-2025
Accepted: 04-05-2025
Page No: 890-911

Abstract
This study determined the conceptions of assessment and assessment practices of the selected private junior high school teachers in the province of Laguna. Based on the results of the study, a proposed enhanced faculty development program was provided.
The researcher utilized a quantitative research design employing an adopted survey questionnaire. One hundred eighty-four (184) private junior high school teachers were randomly selected to serve as respondents. Percentage distribution, mean, and weighted mean were used for the descriptive questions of the study specifically on profile of the respondents, conceptions of assessment and assessment practices. Moreover, Man-Whitney U test and the Kruskal-Wallis H. test were used to determine whether profile and conceptions of assessment have significant difference and whether profile and assessment practices have significant difference.
Results of the study showed that majority of the respondents were female, teaching for 4-6 years, specialized in English, Bachelor's degree holders, classroom teachers, and were able to attend 1-5 trainings related to assessment. Respondents strongly agreed on the conceptions of assessment in terms of school accountability, student accountability, improvement, standards-based, concept development, formative assessment and summative assessment. Additionally, the purpose of assessment and classroom assessment are highly practiced assessment practices than assessment strategies.
Furthermore, there was no significant difference between the conceptions of assessment when respondents were grouped according to sex, number of years in teaching, subject specialization and number of trainings related to assessment attended. There was a significant difference on the conceptions of assessment in terms of school accountability, student accountability, improvement, and formative assessment when respondents were grouped according to highest educational attainment. There was a significant difference observed on the respondents' conceptions of assessment in terms of school accountability and standards-based when they were grouped according to position.
Finally, the study also revealed that there was a significant difference on the respondents' assessment practices in terms of classroom assessment when they were grouped according to sex. But, in terms of purpose of assessment and assessment strategies, there was no significant difference found. There was no significant difference found between the respondents' assessment practices and the number of years in teaching, subject specialization, highest educational attainment and number of trainings about assessment attended. There was significant difference found in terms of purpose of assessment and assessment strategies when the respondents were grouped according to position. But there was no significant difference in terms of classroom assessment when they were grouped based on position.

Keywords: Assessment Practices, Assessment Strategies, Classroom Assessment, Conceptions of Assessment, Faculty Development Program, Formative Assessment, Improvement, Irrelevance, Purpose of Assessment, School Accountability, Student Accountability, Summative Assessment

1. Introduction

Assessment plays an extremely important role in today's educational landscape. Teachers do not just design curriculum and deliver instruction but also play a significant role in the assessment process. They ensure alignment or unity among curriculum, instruction, and assessment.

In this view, Piosang (2017) [39] underscored that assessment serves significant purposes in education. He cited in his study what Ewell (2009) said about assessment where the latter proposed the Assessment Paradigms. Elwell coined the words assessment paradigms when he was referring to the two critical purposes of assessments such as (a) Assessment for improvement paradigm and (b) Assessment for accountability paradigm. He pointed out that assessment is utilized in improving teaching and learning. Furthermore, the second paradigm tells more of assessment as an instrument for measuring accountability of the teachers, schools and even government (2009).

Further, the study of Rural in 2019 revealed that teachers have to remember that if assessment is guided with clear purpose it is appropriate, whether it may be for learning, as learning or of learning and the need for the teachers as assessors to undergo trainings that would give them knowledge and skills in assessing students' learning. The above-mentioned document and the study of Rural showed the call for the educators, schools and most especially teachers to take assessment issue a serious assessment but must acquire the necessary knowledge and skills on how to execute them.

Brown's study in 2004 showed that the respondents agreed on the construct that assessment helps better teaching and learning and holds schools accountable. Nevertheless, the participants disagreed that student accountability and assessment were irrelevant. Meanwhile, the work of Azis (2015) [5] showed that the respondents in the study believed that the purpose of the administration of assessment was to improve teaching and learning and at the same time to hold both schools and students accountable. The respondents also believed that assessment was irrelevant because they gave low ratings on the irrelevant items. Utilizing the same instrument, the results were the same with previous researches made in New Zealand (Brown, 2002), Virginia, USA (Calveric, 2010) Ankara, Turkey, (Vardar, 2010) and the Netherlands (Segers & Tillema, 2011). Also, Barnes *et al*. (2015) discovered that US teachers conceive assessment as a tool for accountability, a way to describing and enhancing teaching and learning yet it is irrelevant. Further, they found that teachers possessed varying and sometimes contradicting conceptions on assessment. Hence, their design and utilization of different assessment practices and their attitude towards learning and development program may be influenced. Lastly, the study of Rural in 2019 disclosed teacher respondents' strong agreement on the construct that assessment is good for school accountability, student accountability and improvement but not with irrelevance- bad and ignore and irrelevance- accurate where the respondents strongly disagree and agree, respectively.

Ballada and Aliño (2018) employed the Brown's COA-III when they explored on Filipino teachers' knowledge and beliefs of assessment. Using exploratory and confirmatory factor analyses, they found out that there is a lack of invariance of the scale. Engagement in learning and development programs as perceived by teachers as a contributing factor to their conceptions and practices came out to be one of the important notes of the study of Azis (2015) [5].

This has inspired the researcher as an instructional leader who likewise holds the belief that assessment is a critical component in our educational system. Specifically, at the ground, he saw the dire need to examine teachers' conceptions of assessment as their beliefs are believed to dictate their classroom practices most especially in terms of classroom assessment.

previously conducted research.

2. Method

The researcher utilized the quantitative approach or method. Quantitative research, according to Yazon (2017) "assumes that phenomena should be studied objectively to obtain a single truth, or at least reality within known probabilities, with an emphasis on measurement, numerical data, and experiments." In this study, the researcher used the quantitative method to quantify the collected data and the analysis as well. Specifically, quantitative research was used in this study to determine teachers' conceptions and practices of assessment.

This study involved select secondary private schools in the province of Laguna (composed of seven divisions: City Schools Divisions of Binan, Cabuyao, Calamba, San Pablo, San Pedro, Santa Rosa and Schools Division of Laguna) as the target population.

Multistage sampling was employed in this study which is an extension of cluster sampling that dealt with the clusters first randomly selected and within the selected clusters, sample units were likewise randomly selected. Furthermore, the study used the stratified proportionate to get the sample population.

The simple random technique, which allowed the researcher to pick random samples from the whole population, was used for this study. It was advantageous for the study because the respondents will not be selected which will ensure that the result will not be subjective.

All in all, there are 184 respondents in this study.

The procedure for this study is as follows:

2.1 Problem formulation

The purpose of this study is to determine teachers' conceptions and practices of assessment.

Specifically, the study sought to answer the following questions:

1. What is the profile of the respondents in terms of:
1.1 sex;
1.2 number of years in teaching;
1.3 subject specialization;
1.4 highest educational attainment;
1.5 position; and
1.6 number of trainings attended related to assessment?

2. What are the teachers' conceptions of assessment in terms of the following:
2.1 school accountability;
2.2 student accountability;
2.3 improvement;
2.4 irrelevance;
2.5 standards-based;
2.6 concept development;
2.7 formative assessment; and
2.8 summative assessment?

3. What is the level of assessment practices among secondary teachers in the province of Laguna in terms of the following:
3.1 purpose of assessment;
3.2 classroom assessment; and
3.3 assessment strategies?

4. Is there a significant difference on the conceptions of the respondents in terms of assessment when they are grouped by profile?
5. Is there a significant

2.2 Literature search

To give a clearer perspective of the study, the researcher conducted a review of related literature and studies. This review focused on the conceptions of assessment and challenges experienced by the teachers in assessment.

Calveric (2010), Ounis (2017) [38], and Linn and Miller (2005) discoursed on assessment as a very important component in the teaching and learning process. They all agree that through assessment educators are helped to measure student learning and enable them to use assessment results to improve instruction and learning.

Ballada and Alino in 2018 mentioned that Kane, Sandretto, and Heath (2002) concluded that the knowledge, beliefs, and thinking of the teachers on the components of the teaching and learning process affect their practices in the classroom. The researchers added that one of those important components is classroom assessment. It is for this reason that teachers' conceptions of assessment need to be examined.

Barnes *et al*. in 2015 mentioned that studying teachers' beliefs and conceptions has an important link with their practices. Brown (2008), Brown *et al*. (2009) and Vineyard and Killen (2007) both agreed the conceptions of teachers dictate their decisions in class and their activities professionally. They believed that different assessment conceptions influence their decisions and professional activities. Hence, this leads to different assessment practices. All these were cited by Monteiro *et al*. in their study in 2021. For instance, if a teacher conceives assessment as important for teaching and learning achievement then he or she will use formative methods of assessment. On the other hand, if he or she sees assessment for accountability then the teacher will utilize summative assessment strategies (Vandeyar and Killen, 2007).

Melaku and Wudu in 2022 concluded that to study conceptions of teachers is essential since it is related to knowledge and beliefs which eventually create impact to teaching practices, including classroom assessment. Varying teachers' conceptions of classroom assessment subsequently may result to different assessment practices. When teachers conceive that assessment is used to improve learning and teaching tend to lead teachers to the practice of formative assessment while those teachers who conceived that students should be responsible for their learning (a conception of assessment for accountability) will tend to favor the formal, summative assessment methods in the classroom (Opre, 2015) [37] and therefore, this kind of teachers do assessment of learning in their classrooms (Monteiro *et al*., 2021) [35]. It is important to note then that the assessment practices of teachers as a product of their conceptions of assessment serve a vital role in the learning progress of the learners.

The Philippine Professional Standards for Teachers which was adopted by DepEd in 2017 assures teacher quality in our educational system from pre-service education to in-service training. As the basis for the teachers development programs, it also deals with assessment specifically Domain 5, known as Assessment and Reporting, is composed of five strands: 1. Design, selection, organization and utilization of assessment strategies, 2. Monitoring and evaluation of learner progress and achievement, 3. Feedback to improve learning, 4. Communication of learner needs, progress, and achievement to key stakeholders, and 5. Use of assessment data to enhance teaching and learning practices and programs.

The above-cited reviewed literature presented how teachers conceived assessment and how they practice assessment while the DepEd designed development programs that focus no longer on the conceptions and therefore cannot provide baseline data in determining whether or not teachers are already equipped with the skills mentioned in the 5 strands under Domain 5 of PPST.

2.3 Data evaluation

The author evaluates the contents of the research journal obtained so that the research data to be discussed can be in accordance with what is desired.

2.4 Data analysis and interpretation

The data collected in this study was organized and classified based on the research design and problems formulated. The data were coded, tallied, and tabulated to facilitate the presentation and interpretation of results using the following:

a) Percentage Distribution. It is a descriptive statistic that was used to determine the number of respondents responding to a category (demographic profile) against the total number of respondents.
b) Mean. The weighted mean was used to calculate the central tendency response of the respondents. Specifically, this was used in determining the respondents' conception of assessment and assessment practices.
c) Mann-Whitney U Test. It is a nonparametric equivalent of independent samples t-test which is commonly used when the data are ordinal. This study utilized a rating scale (1, 2, 3, and 4) which provides ordinal data. Mann-Whitney U test was utilized because the researcher was interested to see whether there was a difference between sex (male and female) in their conceptions of assessment and assessment practices.
d) Kruskal-Wallis's test. The researcher made use of this to conduct a nonparametric analysis and have a single independent measures factor (independent variable) with more than two samples. Specifically, using a rating scale (1, 2, 3 and 4) which provides ordinal data, the researcher was interested to see whether there was a difference between number of years in teaching, subject specialization, position, highest educational attainment and number of trainings attended related to assessment in their conceptions of assessment and assessment practices.

3. Results and Discussion

A. Results

The following are the results of the study.

1. Profile of the Respondents

Table 1: Profile of the Respondents According to Sex

Sex	Frequency	Percentage (%)
Male	51	27.7
Female	133	72.3
Total	**184**	**100**

Table 3 shows that 133 respondents which is equivalent to

72.3% are female while 51 respondents which is equivalent to 27.7% are male.

Table 4 on the next page shows that 77 respondents or 41.80% have been teaching for 1-3 years, 32 respondents which is equal to 17.40% teach for 4-6 years, 24 respondents which is equal to 13% teach for 7-9 years, 15 respondents which is equal to 8.20% teach for 10-12 years, 17 respondents which is equal to 9.20% teach for 13-15 years, 3 respondents which is equal to 1.60% teach for 16-18 years and 16 respondents or 8.70% are teaching for 19 years and above.

Table 2: Profile of the Respondents According to Years in Teaching

Years in Teaching	Frequency	Percentage (%)
1-3 Years	77	41.80
4-6 Years	32	17.40
7-9 Years	24	13.00
10-12 Years	15	8.20
13-15 Years	17	9.20
16-18 Years	3	1.60
19 Years & Above	16	8.70
Total	**184**	**100.00**

Table 3: Profile of the Respondents According to Subject Specialization

Subject Specialization	Frequency	Percentage (%)
Araling Panlipunan	29	15.80
English	34	18.50
Filipino	20	10.90
Mapeh	15	8.20
Math	27	14.70
Science	24	13.00
Technology and Livelihood Education	24	13.00
Edukasyon sa Pagpapakatao	10	5.40
Others	1	0.50
Total	**184**	**100.00**

Table 5 from the previous page reveals that in terms of subject specialization, 29 respondents or 15.8% have specialized in Araling Panlipunan, 34, or 18.5% specialized in English, 20, or 10.9% specialized in Filipino, 15 or 8.2% specialized in MAPEH, 27 or 14.7% specialized in Math, 24 or 13% specialized in Science, 24 or 13% specialized in TLE, 10 or 5.4% specialized in Edukasyon sa Pagpapakatao and 1 or 0.5% has a specialization in other subject.

Table 4: Profile of the Respondents According to Highest Educational Attainment

Highest Educational Attainment	Frequency	Percentage (%)
Doctorate (Grad.)	2	1.10
Doctorate (With Units Earned)	1	0.50
Master's (Grad.)	9	4.90
Master's (With Units Earned)	26	14.10
Bachelor's	146	79.30
Total	**184**	**100.00**

In terms of highest educational attainment, the table above shows that 2 or 1.1% of the respondents obtained a Doctoral degree while 1 or .5% had units in a Doctoral degree. 9 or 4.9% of the respondents were Master's degree holders and 26 or 14.1% of them were able to earn units in their Master's degree. 146 or 79.3% of the respondents obtained a bachelor's degree. The data shows that most of the respondents obtained a bachelor's degree.

Of the 184 respondents, there are 132 or 71.7% of them are classroom teachers. There are 43 or 23.4% coordinators and 7 or 3.8% of them are principals. 2 or 1.1% of the respondents occupy other positions. The majority of the respondents in this study are classroom teachers as the data on the table below shows.

Table 5: Profile of the Respondents According to Position

Position	Frequency	Percentage (%)
Classroom Teacher	132	71.70
Coordinator	43	23.40
Principal	7	3.80
Others	2	1.10
Total	**184**	**100.00**

Table 6: Profile of the Respondents According to Number of Trainings Related To Assessment

No. Of Trainings Related to Assessment	Frequency	Percentage (%)
None	46	25.00
1-5 Times	105	57.10
6-10 Times	18	9.80
11-15 Times	4	2.20
More Than 15 Times	11	6.00
Total	**184**	**100.00**

Table 8 from the previous page reveals that 46 or 25% of the respondents have not attended yet any training related to assessment. There are 105 or 57.1% of them attended 1-5 times, 18 or 9.8% of them attended 6-10 times, 4 04 2.2% attended 11-15 times and 11 or 6% of them attended training more than 15 times.

2. Teachers' conceptions of assessment

Table 7: Respondent's Conceptions of Assessment in Terms of School Accountability

School Accountability	Mean	Verbal Interpretation
1. Assessment provides information on how well schools are doing.	3.58	Strongly Agree
2. Assessment measures the worth or quality of schools.	3.54	Strongly Agree
3. Assessment shows the value schools add to student learning.	3.58	Strongly Agree
4. Assessment keeps schools honest and up to scratch.	3.48	Strongly Agree
Overall Weighted Mean	3.55	Strongly Agree

Table 9 below shows the respondents' conception of assessment in terms of School Accountability. The respondents strongly agreed that assessment can be used to hold schools accountable, with an overall weighted mean of 3.55. As shown on the table, the statements "Assessment provides information on how well schools are doing." and "Assessment shows the value schools add to student learning got the highest mean of 3.58 which was verbally interpreted as strongly agree while the statement "Assessment keeps schools honest and up-to-scratch." obtained the lowest weighted mean of 3.48 but still verbally interpreted as strongly agree.

Table 8: Respondent's Conceptions of Assessment in Terms of Student Accountability

Student Accountability	Mean	Verbal Interpretation
1. Assessment selects students for future education or employment.	3.32	Strongly Agree
2. Assessment is comparing student work against set criteria.	3.37	Strongly Agree
3. Assessment determines if students meet qualifications standards.	3.53	Strongly Agree
4. Assessment is assigning a grade or level to student work.	3.42	Strongly Agree
5. Assessment places students into categories.	3.37	Strongly Agree
6. Assessment is checking off progress against achievement objectives.	3.47	Strongly Agree
7. Assessment is completing checklists.	3.33	Strongly Agree
Overall Weighted Mean	3.40	Strongly Agree

Results showed that junior high school teachers find school accountability as a very important factor in assessment. This necessitates that the school should make sure that teachers possess full understanding what assessment stands for and the implication it creates towards the school performance.

Given the data above, Piosang (2017) [39] cited Ewell in 2009 argued that one of the two critical purposes of assessments is the accountability of the teachers and school heads (school) for student achievement.

Table 10 above shows the strong agreement among respondents towards student accountability, with an overall weighted mean of 3.40. The statements: "Assessment determines if students meet qualifications standards.", "Assessment is checking off progress against achievement objectives" and "Assessment is assigning a grade or level to student work." got the highest weighted mean of 3.53, 3.47, and 3.42, respectively, and all verbally interpreted as strongly agree.

Meanwhile, "Assessment selects students for future education or employment" obtained the lowest weighted mean of 3.32 and was still interpreted as strongly agree. The result showed that junior high school teachers also find student accountability as a strong factor in assessment. This implies that the teachers must be equipped with the skills on how to ensure that assessment is being used to prepare the learners to their future education or employment.

This finding is of importance because if teachers are aware of this then they will make sure that the assessment they would give to learners will provide the latter with the opportunity to involve them in the process of assessment by reflecting on the results and finding ways on how to improve their performance. The teachers make sure that they are also able to communicate the purpose of the assessments they administer. Teachers should ensure that they provide learners opportunities to self-reflect using assessment results.

The finding is likewise consistent to what purpose classroom assessment should serve for and that is to promote self-reflection and personal accountability among students about their own learning (DepEd, 2015) [20].

Table 9: Respondent's Conceptions of Assessment in Terms of Improvement

Improvement	Overall Weighted Mean	Verbal Interpretation
1. Improvement-Describe	3.56	Strongly Agree
2. Improvement-Student Learning	3.49	Strongly Agree
3. Improvement-Validity	3.19	Agree
4. Improvement-Teaching	3.47	Strongly Agree
Average Overall Weighted Mean	3.42	Strongly Agree

Table 11 on the next page shows the overall weighted mean and verbal interpretation of respondents' conceptions of assessment in terms of improvement. With an average overall weighted mean of 3.42, teachers expressed strong agreement on the construct that assessment serves the function of improvement. The conception that while assessment improves it also describes the performance of the learners got the highest overall weighted mean of 3.56 verbally interpreted as strongly agree.

On the other hand, the belief that assessment results to improvement and that it is valid got the lowest overall weighted mean of 3.19 verbally interpreted as agree.

Apparently, teachers conceived assessment as a tool to help improve student learning and in the process feedback plays significant role. It is in this note that the study of Magno and Ocampo (2018) [29] is worth mentioning. The focus of the said study was on the value for Formative Assessment in which the researchers cited the work of Deci and Ryan (2008) that used the Self-Determination Theory and opined that feedback serves as external regulation to improve on the learner's performance. This shows the strong connection between feedback and learning improvement.

This result echoes what Black and William mentioned in 2018 that assessment supports learning when learners receive feedback that takes learning forward.

Teachers must be given trainings on the construction of a valid assessment so that the results will be trustworthy, dependable and can predict the future learners performance. Additionally, teachers should keep in mind that assessment should be objective and results are consistent.

The result also supported the study conducted by Barnes *et al*., (2017) [9] which revealed that US teachers have three dominant conceptions of assessment and one of which is the conception that assessment is a tool for improving teaching and learning as well.

Table 10: Respondent's Conceptions of Assessment in Terms of Irrelevance

Irrelevance	Overall Weighted Mean	Verbal Interpretation
1. Irrelevance-Bad	2.23	Disagree
2. Irrelevance-Ignored	2.03	Disagree
3. Irrelevance-Accurate	2.79	Agree
Average Overall Weighted Mean	2.35	Disagree

Table 12 shows the average overall weighted mean and verbal interpretation of respondents' conceptions of assessment in terms of irrelevance. With the average overall weighted mean of 2.35, the teachers disagreed that assessment is irrelevant. However, the respondents agreed to the conception that assessment is irrelevant because it provides little useful information due to inaccuracy or lack of precision as it got the highest overall weighted mean that is equivalent to 2.79. Believing that assessment is irrelevant because it is conducted or administered yet it is ignored obtained the lowest overall weighted mean that is equivalent to 2.03. This means that the respondents disagreed with the said statement.

This finding implies that the respondents did not consider assessment as a bad factor. The negative conception that assessment is not a good practice was disapproved by the respondents.

Likewise, this suggests that respondents agreed that accuracy is also an important factor to consider in assessment. Teachers still believed that assessment must be accurate so that results will be valid, reliable, and useful for both teachers and learners.

Table 11: Respondent's Conceptions of Assessment in Terms of Standards-Based

Standards-Based	Mean	Verbal Interpretation
1. Assessment helps identify and set the essential knowledge and understanding that should be learned.	3.52	Strongly Agree
2. Assessment helps cover a specified scope of sequential topics (spiral progressions) within each learning strand, domain, theme or component.	3.51	Strongly Agree
3. Assessment answers the question, "What should the learners know?".	3.47	Strongly Agree
4. Assessment helps develop the abilities and skills that learners are expected to demonstrate about the content standards and integration of 21st century skills.	3.52	Strongly Agree
5. Assessment guides learners to apply what they know.	3.52	Strongly Agree
6. Assessment determines how well must learners do their work.	3.48	Strongly Agree
7. Assessment identifies how learners use their learning or understanding in different situations.	3.54	Strongly Agree
8. Assessment provides tools in measuring how learners demonstrate what they know.	3.52	Strongly Agree
9. Assessment determines learners' knowledge, skills, and attitudes they need to demonstrate in every lesson and/or learning activity.	3.55	Strongly Agree
Overall Weighted Mean	3.51	Strongly Agree

The results from Table 13 show that the respondents strongly agreed on the conceptions of assessment in terms of standards-based with the overall weighted mean of 3.51. The statement "Assessment determines learners' knowledge, skills and attitudes they need to demonstrate in every lesson and/or learning activity" obtained the highest weighted mean of 3.55, verbally interpreted as strongly agree while the third statement "Assessment answers the question, "What should the learners know?" got the lowest weighted mean of 3.47 which was also verbally interpreted as strongly agree.

Respondents understood that in designing assessment they understand that assessment is being carried out measuring learners' performance in relation to learning standards. By learning standards, these comprised the content standards, performance standards and learning competencies (DepEd, 2015). This implies that teachers make sure design and conduct assessment that is standards-based. In other words, teachers measure what is to be measured.

This is a confirmation to the adopted Philippine Professional Standards for Teachers where teachers are expected to be skilled in the designing, selecting, organizing, and utilizing of assessment strategies. Hence, this ensures that the assessment given to learners has a basis, that is the standards.

Table 12: Respondent's Conceptions of Assessment in Terms of Concept Development

Concept Development	Mean	Verbal Interpretation
1. The learner can recall information and retrieve relevant knowledge from long-term memory.	3.32	Strongly Agree
2. The learner can generate meaning from oral, written and graphic messages.	3.42	Strongly Agree
3. The learner can distinguish between parts and determine how they relate to one another and to the overall structure and purpose.	3.38	Strongly Agree
4. The learner can use information to undertake a procedure in familiar situation or in a new way.	3.41	Strongly Agree
5. The learner can make judgments and justify decision.	3.35	Strongly Agree
6. The learner can put elements together to form a functional whole and create a new product or point of view.	3.37	Strongly Agree
Overall Weighted Mean	3.38	Strongly Agree

Table 14 on the succeeding page presents mean and verbal interpretations of respondents' conceptions on assessment in terms of concept development. Of the six statements, the statement "The learner can generate meaning from oral, written and graphic messages" got the highest weighted mean of 3.42 which was verbally interpreted as strongly agree. On the other hand, the statement "The learner can recall information and retrieve relevant knowledge from long-term memory" obtained the lowest weighted mean which is 3.32 but still verbally interpreted as strongly agree.

This implicates that concept development is also one of the considerations in doing assessment. This is an affirmation that as teachers formulate assessment tasks and activities, they have to make sure that knowledge/concepts and skills are developed from the simplest to the complex ones. Such progression in the development is best operationalized using Anderson and Krathwohl's (2001) Cognitive Process Dimension (DepEd, 2015) [20].

Table 13: Respondent's Conceptions of Assessment in Terms of Formative Assessment

Formative Assessment	Mean	Verbal Interpretation
1. Formative assessment is ongoing and linked to the learning process.	3.43	Strongly Agree
2. Formative assessment is informal.	2.95	Agree
3. Formative assessment provides immediate feedback to both learner and teacher.	3.42	Strongly Agree
4. Formative assessment helps prepare learners for summative assessments.	3.45	Strongly Agree
5. Formative assessment is recorded but not included as a basis for grading.	3.19	Agree
6. Formative assessment may be done before, during and after the lesson.	3.29	Strongly Agree
7. Formative assessment identifies what hinders learning.	3.33	Strongly Agree
8. Formative assessment identifies learning gaps.	3.41	Strongly Agree
9. Formative assessment tracks learner progress in comparison to Formative assessment results prior to the lesson proper.	3.40	Strongly Agree
10. Formative assessment helps make decisions on whether to proceed with the next lesson, re-teach or provide for corrective measures or reinforcements.	3.39	Strongly Agree
Overall Weighted Mean	3.33	Strongly Agree

Table 15 shows that respondents strongly agreed with all statements except for the statements "Formative assessment is recorded but not included as a basis for grading" and "Formative assessment is informal" which got the weighted means of 3.19 and 2.95, respectively, both verbally interpreted as agree. Nonetheless, the overall weighted mean of the respondents' conception of assessment in terms of Formative Assessment is equal to 3.33 which received strong agreement from the respondents.

This implies that teachers understand that formative assessment is an important process that could help them see how they could be of help to their learners since the process could give the chance to assess what appropriate intervention and assistance are necessary to be done and how much.

Considering the findings, Bernardo (2018) emphasized the need for both teachers and learners to have a clear understanding on the learning intentions and that they are also aware of where the learners are relative to these learning intentions so they can work together to more effectively move forward to their learning intentions.

Given this, Bernardo (2018) made a remarkable note in his work that teachers can take advantage of what constitutes appropriate and useful feedback mentioning that this makes up a very crucial component of formative assessment. He

cited some pointers for giving effective feedback as suggested by Sutton in 1998. One of which is that for feedback to be effective, it should be very specific. Instead of using comments such as "very good' "good job" "wrong!" or "do better next time" which are evaluative in nature teachers should make use of descriptive feedback. While evaluative feedback involves some value judgment (Tunstall & Gipps, 1996) descriptive feedback involves comments on the student's work like "Your work is good because you explained the idea through a unique or practical example." The latter provides clearer feedback on his work and therefore learners will know how they can meet what is expected of them. To be useful, Sutton reiterated that specific and descriptive feedback should be given to the learners at the soonest possible time- immediate feedback.

Table 14: Respondent's Conceptions of Assessment in Terms of Summative Assessment

Summative Assessment	Mean	Verbal Interpretation
1. Summative assessment is an assessment of learning given at the end of a particular unit.	3.61	Strongly Agree
2. Summative assessment usually occurs towards the end of a period of learning in order to describe the standards reached by the learner.	3.56	Strongly Agree
3. Summative assessment is usually for the benefit of people other than the learner.	2.96	Agree
4. Summative assessment results are recorded and are included in the computation of the final grade.	3.53	Strongly Agree
5. Written works as a form of Summative assessment ensures that students are able to express learned skills and concepts in written form.	3.49	Strongly Agree
6. Written works as a form of Summative assessment includes essays, written report, long quizzes and other written outputs.	3.44	Strongly Agree
7. The items in long quizzes or tests as a form of Summative assessment are distributed across the Cognitive Process Dimensions.	3.43	Strongly Agree
8. Written works help practice and prepare for quarterly assessment and other standardized assessments.	3.49	Strongly Agree
9. Performance tasks involve students in the learning process individually or in collaboration with teammates over a period of time.	3.55	Strongly Agree
10. Performance tasks give students opportunities to demonstrate and integrate their knowledge, understanding, and skills about topics or lessons learned in a specific real-life situation.	3.59	Strongly Agree
11. Performance tasks give students the freedom to express their learning in appropriate and diverse ways.	3.60	Strongly Agree
12. Performance tasks encourage student inquiry, and integration of knowledge, understanding, and skills in various contexts beyond the assessment period.	3.59	Strongly Agree
13. Quarterly assessments synthesize all the learning skills, concepts, and values learned in an entire quarter.	3.54	Strongly Agree
14. Quarterly assessments measure student learning at the end of the quarter.	3.57	Strongly Agree
15. Quarterly assessments are in the form of objective tests, performance-based assessments, or a combination thereof.	3.57	Strongly Agree
Overall Weighted Mean	3.50	Strongly Agree

It can be seen from Table 16 from the previous page that the respondents had a strong agreement on the conceptions of assessment in terms of summative assessment with an overall weighted mean of 3.50. All 14 items obtained a strongly agree rating from the respondents and only the statement "Summative assessment is usually for the benefit of people other than the learner" got an agree response with a weighted mean of 2.96. The first statement "Summative assessment is an assessment of learning given at the end of a particular unit" got the highest weighted mean of 3.61 verbally interpreted as strongly agree.

The result is consistent with what DepEd Order No. 8, s. 2015 emphasized that the summative assessment is given towards the end of a particular unit and is usually given to learners to know where they are in relation to the learning standards.

3. Respondent's level of assessment practices

Table 15: Respondent's Level of Assessment Practices in Terms of Purpose of Assessment

Purpose of Assessment	Mean	Verbal Interpretation
1. I guide students to set their goals and monitor their own learning progress.	3.64	Highly Practiced
2. I demonstrate to students how to do self-assessment.	3.49	Highly Practiced
3. I determine how students can learn on their own in class.	3.61	Highly Practiced
4. I assist students to identify means of getting personal feedback and monitoring their own learning process.	3.58	Highly Practiced
5. I help students develop clear criteria of a good learning practice.	3.55	Highly Practiced
6. I set the criteria for students to assess their own performance in class.	3.57	Highly Practiced
7. I measure the extent of learning at the end of a lesson or subject.	3.60	Highly Practiced
8. I evaluate the level of competence of students at the end of an instructional program.	3.60	Highly Practiced
9. I determine the degree of accomplishment of a desired learning outcome at the end of	3.55	Highly Practiced

a lesson.		
10. I make the final decision about the level of learning that students achieved at the end of a lesson or subject.	3.47	Highly Practiced
11. I rank the students based on their class performance to inform other school officials.	3.23	Practiced
12. I provide information to parents about the performance of their children in school.	3.57	Highly Practiced
13. I examine how one student performs compared to others in my class.	3.36	Highly Practiced
14. I provide information to other teachers, schools and employers regarding students' performance in class.	3.42	Highly Practiced
15. I help students improve their learning process and class performance.	3.64	Highly Practiced
16. I assist students to determine their learning strengths and weaknesses in class.	3.61	Highly Practiced
17. I identify better learning opportunities for students in class.	3.64	Highly Practiced
18. I periodically collect learning data from students to improve the instructional process.	3.52	Highly Practiced
Purpose of Assessment	Mean	Verbal Interpretation
19. I ask questions or tasks that allow me to know whether students:		
a. can recall or remember what is taught in class	3.66	Highly Practiced
b. explain Ideas and concepts	3.66	Highly Practiced
c. use learned information or concepts in new way	3.63	Highly Practiced
d. analyze a situation or condition	3.62	Highly Practiced
e. justify a stand or decision	3.61	Highly Practiced
f. create a new product or point of view or idea	3.63	Highly Practiced
Overall Weighted Mean	3.56	Highly Practiced

Table 17 on the succeeding pages demonstrates mean and verbal interpretation of respondents' assessment practices in terms of purpose of assessment. Out of 24 statements, 23 items were verbally interpreted as highly practiced by the respondents except for the statement "I rank the students based on their class performance to inform other school officials" which got the lowest weighted mean of 3.23 verbally interpreted as practiced. The statements "I ask questions or tasks that allow me to know whether students: can recall or remember what is taught in class" and "…explain Ideas and concepts" are highly practiced since both received the highest weighted mean of 3.66. This area received an overall weighted mean of 3.56, verbally interpreted as highly practiced.

The data are a confirmation of what Magno and Ocampo (2018) [29] explained that learners need to be informed that classroom assessment including other activities must be consistent with the learning targets set by the teachers, students, curriculum, and other external standards. They expounded that once learners are conscious of learning targets, they are aware of the goals that need to be attained while engaging in the assigned tasks.

The results imply that teachers recognized that the indicators are very important to help them communicate to learners, parents, teachers and other members of the school the purpose of assessment and subsequently, aid the learners focused in doing the task).

Table 16: Respondent's Level of Assessment Practices in Terms of Classroom Assessment

Classroom Assessment	**Mean**	**Verbal Interpretation**
1. I prepare at least 3 learning objectives.	3.58	Highly Practiced
2. I refer to the curriculum when I organize my learning objectives.	3.65	Highly Practiced
Classroom Assessment	**Mean**	**Verbal Interpretation**
3. I follow taxonomy in preparing learning objectives.	3.51	Highly Practiced
4. I prepare a test plan according to the learning of my lessons.	3.66	Highly Practiced
5. I ensure that every topic I cover in class is included in the assessment plan.	3.65	Highly Practiced
6. I relate to the instructional process with the assessment process.	3.60	Highly Practiced
7. I try to include a variety of questions to measure different levels of cognitive skills.	3.62	Highly Practiced
8. I ensure that appropriate assessment strategies are employed.	3.63	Highly Practiced
9. I prepare table of specifications (TOS).	3.34	Highly Practiced
10. I write clear learning objectives so that students are aware of what is to be assessed.	3.62	Highly Practiced
11. I use textbooks as references when I write test items.	3.50	Highly Practiced
12. I include a variety of questions in a single test.	3.58	Highly Practiced
13. I make sure I give clear instructions for every type of question I include in a test.	3.70	Highly Practiced
14. I arrange test questions from easy to difficult.	3.58	Highly Practiced
15. I ensure that questions and options are on the same page.	3.63	Highly Practiced
16. I avoid including items that suggest racial, ethnic or gender biases.	3.59	Highly Practiced
17. I try to prepare questions that minimize guessing.	3.60	Highly Practiced
18. I explain the basis of scoring problems solving items to students.	3.58	Highly Practiced
19. I include on the same page the diagrams or maps needed in a particular question.	3.58	Highly Practiced
20. I proofread all test questions and instructions before printing them.	3.67	Highly Practiced
21. I ensure that the classroom is conducive for testing activities.	3.67	Highly Practiced
22. I see to it that cheating is not encouraged in the classroom.	3.72	Highly Practiced
23. I prepare scoring criteria or rubrics before I start making test papers.	3.66	Highly Practiced
24. I score test papers at random	3.21	Practiced
25. I ensure that I have enough test materials before I administer a test.	3.67	Highly Practiced
26. I follow scoring criteria strictly when marking test papers.	3.64	Highly Practiced

Classroom Assessment	Mean	Verbal Interpretation
27. I make sure I have enough time to score test papers.	3.67	Highly Practiced
28. I provide feedback to students after every test.	3.57	Highly Practiced
29. I give a grade equivalent to the total score in a test.	3.55	Highly Practiced
30. I explain to the students how scores are derived.	3.64	Highly Practiced
31. I share test results with other teachers and the school director if necessary.	3.48	Highly Practiced
32. I make sure parents are informed of the test results of their children.	3.59	Highly Practiced
33. I determine the difficulty level of each test item after a test.	3.53	Highly Practiced
34. I conduct an item analysis to know whether items can discriminate students' abilities.	3.35	Highly Practiced
35. I make simple item banking for every subject.	3.36	Highly Practiced
36. I post the names of students who performed well in a test.	3.10	Practiced
37. I return all marked test papers to students on time.	3.59	Highly Practiced
Overall Weighted Mean	3.56	Highly Practiced

Shown in Table 18 are mean and verbal interpretations of respondents' assessment practices in terms of classroom assessment. With the exemption of statements "I post the names of students who performed well in a test" and "I score the papers at random." which got the lowest two weighted means of 3.10 and 3.21, both verbally interpreted as practiced, the 35 statements were conceived to be highly practiced by the respondents. With a weighted mean of 3.72, verbally interpreted as highly practiced, the statement "I see to it that cheating is not encouraged in the classroom" got the highest weighted mean. As the result showed, this area received an overall weighted mean of 3.56 which was verbally interpreted as highly practiced.

This implies that classroom assessment is a non-negotiable process that is designed and administered by teachers to their students. Right conception on why classroom assessment is given, how it is administered, and how results are used will guide teachers in their performance as assessors of student learning.

Table 17: Respondent's Level of Assessment Practices in Terms of Assessment Strategies

Assessment Strategies	Mean	Verbal Interpretation
I use the following assessment strategies:		
a. Multiple choice	3.58	Highly Practiced
b. True-False or Right or wrong	3.54	Highly Practiced
c. Matching types	3.36	Highly Practiced
d. Fill in the blanks or short constructed response	3.31	Highly Practiced
e. Word problems	3.26	Highly Practiced
f. Performance assessment	3.58	Highly Practiced
g. Portfolio Assessment	3.16	Practiced
h. Graded Recitation	3.55	Highly Practiced
i. Observations	3.40	Highly Practiced
j. Term papers or projects	3.37	Highly Practiced
k. Class presentations	3.53	Highly Practiced
l. Interviews and conferences	3.07	Practiced
m. student reflection/ journal writing/ student self-assessment	3.35	Highly Practiced
n. assignments	3.48	Highly Practiced
Overall Weighted Mean	3.40	Highly Practiced

It can be observed from Table 19 that in terms of Assessment Strategies, the overall weighted mean is 3.40 verbally interpreted as highly practiced. Two strategies "Portfolio Assessment" and "Interviews and conferences" were got the weighted means of 3.16 and 3.07, respectively, and verbally interpreted as practiced. On the contrary, there are 12 strategies from the table that were highly practiced by the respondents.

Further, the assessment strategy "Multiple Choice" got the highest weighted mean of 3.58 which was verbally interpreted as highly practiced while "Interviews and conferences" obtained the lowest weighted mean of 3.07, verbally interpreted as practiced.

This implies that the schools must provide mentoring or training that focuses on the development of teachers' skills on how to employ interviews and conferences and other not so popular assessment strategies. Teachers still need to be trained on how they can also design and utilize other assessment strategies that develop higher-order skills.

4. Significant difference on the conceptions of the respondents in terms of assessment when they are grouped by profile

Table 18: Summary Table Showing the Significant Difference on the Respondent's Conception of Assessment When Grouped According to Sex

Conception of Assessment (COA-III, Brown, 2004 & Rural, 2019)	Sex	Mean Rank	Mann-Whitney U	p-value	Decision	Remark
1. School Accountability	Male	84.66	2991.50	0.162	Fail to Reject Ho	Not Significant
	Female	95.51				
2. Student Accountability	Male	88.13	3168.50	0.434	Fail to Reject Ho	Not Significant
	Female	94.18				
3. Improvement	Male	91.64	3240.75	0.610	Fail to Reject Ho	Not Significant
	Female	92.83				
4. Irrelevance	Male	96.79	3172.83	0.477	Fail to Reject Ho	Not Significant
	Female	90.85				
5. Standards-Based	Male	89.78	3253.00	0.621	Fail to Reject Ho	Not Significant
	Female	93.54				
6. Concept Development	Male	99.10	3055.00	0.251	Fail to Reject Ho	Not Significant
	Female	89.97				
7. Formative Assessment	Male	89.32	3229.50	0.580	Fail to Reject Ho	Not Significant
	Female	93.72				
8. Summative Assessment	Male	101.48	2933.50	0.096	Fail to Reject Ho	Not Significant

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 20 on the next page shows the Mann-Whitney U test to verify if there exists a significant difference on the conceptions of the respondents on assessment in terms of sex. It reveals that overall, there is no significant difference between the respondents' conception of assessment and their sex since all the p values are greater than the 0.05 level of significance. The factor standards-based got the highest p-value of 0.621 which is very much greater than the 0.05 level of significance. This is why the null hypothesis was accepted.

On the other side, the factor summative assessment obtained a p-value of 0.096 which is the lowest but is still greater than the 0.05 level of significance. Hence, the null hypothesis was accepted.

Based on a female-dominated respondents survey used in this study, with 133 or 72.3% female respondents, the data showed that the maleness and femaleness of a teacher is not a factor that could influence how a teacher conceives assessment.

Table 19: Summary Table Showing the Significant Difference on the Respondent's Conception of Assessment When Grouped According to Number of years in teaching

Conception of Assessment (COA-III, Brown, 2004 & Rural, 2019)	Number of years in teaching	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. School Accountability	1-3 Years	99.83	9.050	0.171	Fail to Reject Ho	Not Significant
	4-6 Years	92.73				
	7-9 Years	80.40				
	10-12 Years	83.97				
	13-15 Years	104.47				
	16-18 Years	98.67				
	19 Years & Above	69.03				
2. Student Accountability	1-3 Years	94.97	4.400	0.623	Fail to Reject Ho	Not Significant
	4-6 Years	94.69				
	7-9 Years	79.56				
	10-12 Years	89.97				
	13-15 Years	104.06				
	16-18 Years	111.00				
	19 Years & Above	82.28				
3. Improvement	1-3 Years	124.08	5.531	0.543	Fail to Reject Ho	Not Significant
	4-6 Years	87.31				
	7-9 Years	81.42				
	10-12 Years	91.63				
	13-15 Years	91.48				
	16-18 Years	95.17				
	19 Years & Above	89.21				
4. Irrelevance	1-3 Years	99.13	7.852	0.324	Fail to Reject Ho	Not Significant
	4-6 Years	81.90				
	7-9 Years	80.84				
	10-12 Years	94.75				
	13-15 Years	84.53				
	16-18 Years	88.89				
	19 Years & Above	107.00				
5. Standards-Based	1-3 Years	97.95	6.022	0.421	Fail to Reject	Not

	4-6 Years	92.73			Ho	Significant
	7-9 Years	74.00				
	10-12 Years	89.70				
	13-15 Years	99.76				
	16-18 Years	72.00				
	19 Years & Above	92.28				
6. Concept Development	1-3 Years	101.21	10.458	0.107	Fail to Reject Ho	Not Significant
	4-6 Years	84.72				
	7-9 Years	78.10				
	10-12 Years	114.17				
	13-15 Years	83.00				
	16-18 Years	83.83				
	19 Years & Above	79.16				
7. Formative Assessment	1-3 Years	101.31	10.797	0.095	Fail to Reject Ho	Not Significant
	4-6 Years	83.16				
	7-9 Years	75.33				
	10-12 Years	106.33				
	13-15 Years	100.38				
	16-18 Years	85.17				
	19 Years & Above	74.59				
8. Summative Assessment	1-3 Years	97.71	12.212	0.057	Fail to Reject Ho	Not Significant
	4-6 Years	86.95				
	7-9 Years	84.56				
	10-12 Years	107.73				
	13-15 Years	106.91				
	16-18 Years	69.17				
	19 Years & Above	65.22				

Note*:* If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 21 on the following pages presents the result when Kruskal-Wallis H was utilized to verify if there exists a significant difference on the respondents' conception of assessment when they were grouped according to number of years in teaching. Apparently, there is no significant difference between their conceptions of assessment and the number of years they are teaching since the p-values are all greater than the 0.05 level of significance and therefore the null hypothesis is accepted.

The factor student accountability obtained the highest p-value which is 0.623 and is greater than the 0.05 level of significance therefore the null hypothesis is accepted. This implies that regardless of the years in teaching, it does not affect one's conception that assessment is indeed significant in aiding students learning. On one hand, the factor summative assessment got the lowest p-value of 0.057 which is also greater than the 0.05 level of significance and therefore the null hypothesis is accepted. Although this means there is no significant difference between the conceptions of assessment and the number of years teaching but the p-value is almost near to the level of significance which means that the number of years teaching where teachers are exposed to and doing it can also affect one's conception of summative assessment.

This result is similar to what Balagtas, Dacanay, Dizon, and Duque (2010) revealed, as cited by Ballada and Aliño in 2018, in a study that examined the literacy of Filipino pre-service and in-service teachers. The mentioned researchers explained that using a self-report survey, they found out that Filipino teachers both pre-service and in-service lack the skills concerning the following competencies: administering scoring, and interpreting results of externally -produced and teacher-made test, using assessment results to make different types of academic decision to students, parents and other stakeholders and recognizing unethical, illegal and other inappropriate methods and use of assessment. The status of "pre-service" is understood as having no experience of real teaching yet while "in-service" denotes that regardless of years in teaching, the number of years of actual teaching as the data above shows does not affect teachers' conceptions on assessment.

However, the results above are inconsistent with what Avvisati (2018) [6] reported that those schools with more experienced teachers tend to perform better in PISA.

Table 22: Summary Table Showing the Significant Difference on the Respondent's Conception of Assessment When Grouped According to Subject specialization

Conception of Assessment (COA-III, Brown, 2004 & Rural, s. 2019)	Subject specialization	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. School Accountability	AP	86.91	3.121	0.874	Fail to Reject Ho	Not Significant
	English	95.15				
	Filipino	83.98				
	MAPEH	83.50				
	Math	90.93				
	Science	92.90				
	TLE	102.23				
	EsP	101.05				
2. Student Accountability	AP	84.62	6.007	0.539	Fail to Reject Ho	Not Significant
	English	104.40				

	Filipino	84.40				
	MAPEH	80.83				
	Math	95.30				
	Science	86.25				
	TLE	100.92				
	EsP	86.70				
3. Improvement	AP	80.78	5.85	0.569	Fail to Reject Ho	Not Significant
	English	92.92				
	Filipino	85.25				
	MAPEH	90.40				
	Math	97.82				
	Science	95.42				
	TLE	96.73				
	EsP	102.06				
4. Irrelevance	AP	95.38	9.377	0.372	Fail to Reject Ho	Not Significant
	English	84.33				
	Filipino	111.37				
	MAPEH	101.05				
	Math	80.96				
	Science	97.77				
	TLE	97.79				
	EsP	73.03				
5. Standards-Based	AP	79.28	7.196	0.409	Fail to Reject Ho	Not Significant
	English	85.31				
	Filipino	91.43				
	MAPEH	83.57				
	Math	99.80				
	Science	97.50				
	TLE	106.96				
	EsP	95.30				
6. Concept Development	AP	71.29	9.139	0.243	Fail to Reject Ho	Not Significant
	English	92.93				
	Filipino	86.90				
	MAPEH	89.30				
	Math	105.13				
	Science	94.48				
	TLE	98.02				
	EsP	107.30				
7. Formative Assessment	AP	77.29	7.711	0.359	Fail to Reject Ho	Not Significant
	English	103.28				
	Filipino	81.40				
	MAPEH	90.77				
	Math	94.43				
	Science	93.92				
	TLE	89.92				
	EsP	113.20				
8. Summative Assessment	AP	78.88	11.954	0.102	Fail to Reject Ho	Not Significant
	English	89.69				
	Filipino	88.35				
	MAPEH	74.30				
	Math	113.61				
	Science	96.10				
	TLE	98.33				
	EsP	88.35				

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 22 from the previous page reflects that in terms of the conception of assessment and the subject specialization, there is no significant difference in all aspects. While school accountability got the highest p-value which is equivalent to 0.874, summative assessment obtained the lowest value which is 0.102. Since both values were greater than the level of significance, which is 0.05, null hypothesis is accepted. Therefore, subject specialization does not influence conceptions of assessment.

Table 23: Summary Table Showing the Significant Difference on the Respondent's Conception of Assessment When Grouped According to Highest Educational Attainment

Conception of Assessment (COA-III, Brown, 2004 & Rural, 2019)	Highest educational attainment	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. School Accountability	Master's (Grad.)	51.00	11.285	0.004	Reject Ho	Significant
	Master's (With Units Earned)	75.96				
	Bachelor's	96.14				
2. Student Accountability	Master's (Grad.)	55.61	6.668	0.036	Reject Ho	Significant
	Master's (With Units Earned)	84.04				
	Bachelor's	94.42				
3. Improvement	Master's (Grad.)	77.96	7.677	0.040	Reject Ho	Significant
	Master's (With Units Earned)	77.25				
	Bachelor's	95.42				
4. Irrelevance	Master's (Grad.)	86.52	4.104	0.230	Fail to Reject Ho	Not Significant
	Master's (With Units Earned)	74.88				
	Bachelor's	94.14				
5. Standards-Based	Master's (Grad.)	89.61	5.599	0.061	Fail to Reject Ho	Not Significant
	Master's (With Units Earned)	71.77				
	Bachelor's	94.51				
6. Concept Development	Master's (Grad.)	102.22	3.171	0.205	Fail to Reject Ho	Not Significant
	Master's (With Units Earned)	76.44				
	Bachelor's	92.90				
7. Formative Assessment	Master's (Grad.)	88.28	9.066	0.011	Reject Ho	Significant
	Master's (With Units Earned)	65.38				
	Bachelor's	95.73				
8. Summative Assessment	Master's (Grad.)	92.67	3.113	0.211	Fail to Reject Ho	Not Significant

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 23 shows the Mann-Whitney U test to verify if there exists a significant difference on the conceptions of the respondents on assessment in terms of highest educational attainment. It reveals that overall, there is no significant difference between the respondents' conception of assessment and their position except school accountability, student accountability, improvement and formative assessment where a significant difference was observed. This implies that the training program to be given to teachers should take into account their highest educational background.

Table 24: Summary Table Showing the Significant Difference on the Respondent's Conception of Assessment When Grouped According to Position

Conception of Assessment (COA-III, Brown, 2004 & Rural, 2019)	Position	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. School Accountability	Classroom Teacher	96.23	6.657	0.036	Reject Ho	Significant
	Coordinator	82.47				
	Principal	57.86				
2. Student Accountability	Classroom Teacher	92.88	1.636	0.441	Fail to Reject Ho	Not Significant
	Coordinator	90.78				
	Principal	69.93				
3. Improvement	Classroom Teacher	95.33	4.570	0.128	Fail to Reject Ho	Not Significant
	Coordinator	84.11				
	Principal	67.79				
4. Irrelevance	Classroom Teacher	94.91	2.90	0.337	Fail to Reject Ho	Not Significant
	Coordinator	81.46				
	Principal	88.95				
5. Standards-Based	Classroom Teacher	93.21	6.249	0.044	Reject Ho	Significant
	Coordinator	93.13				
	Principal	49.29				
6. Concept Development	Classroom Teacher	91.87	0.540	0.764	Fail to Reject Ho	Not Significant
	Coordinator	92.47				
	Principal	78.57				
7. Formative Assessment	Classroom Teacher	94.62	2.451	0.294	Fail to Reject Ho	Not Significant
	Coordinator	84.97				

	Principal	72.79				
8. Summative Assessment	Classroom Teacher	94.95	3.305	0.192	Fail to Reject Ho	Not Significant
	Coordinator	84.10				
	Principal	71.79				

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 24 shows the Kruskal-Wallis H test to verify if there exists a Significant Difference on the conceptions of the respondents on assessment in terms of Position. It
reveals that overall, there is no significant difference between the respondents' conception of assessment and their position except for school accountability and standards-based factors where a significant difference was observed.
Conceptions of assessment in terms of school accountability and Standards-based were observed to vary depending on their positions (Classroom teacher, Coordinator and Principal). This means that a classroom teacher, coordinator and principal may have contradicting belief in seeing assessment to hold school accountable and in conceiving that assessment must be based on standards.
In view of the above results, Philippine Professional Standards for Teachers which was adopted and implemented by the Department of Education since 2017 through DepEd Order No. 42, s. 2017 reiterated the need for every teacher to take into consideration Domain 5 of the standards which is known as Assessment and Reporting, is composed of five strands: 1. Design, selection, organization and utilization of assessment strategies, 2. Monitoring and evaluation of learner progress and achievement, 3. Feedback to improve learning, 4. Communication of learner needs, progress and achievement to key stakeholders, and 5. Use of assessment data to enhance teaching and learning practices and programs.

Table 25: Summary Table Showing the Significant Difference on the Respondent's Conception of Assessment When Grouped According to Number of trainings attended related to Assessment

Conception of Assessment (COA-III, Brown, 2004 & Rural, s. 2019)	Number of trainings attended related to Assessment	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. School Accountability	None	90.05	6.040	0.196	Fail to Reject Ho	Not Significant
	1-5 Times	97.66				
	6-10 Times	81.17				
	11-15 Times	106.75				
	More Than 15 Times	66.86				
2. Student Accountability	None	98.00	2.985	0.560	Fail to Reject Ho	Not Significant
	1-5 Times	91.46				
	6-10 Times	81.50				
	11-15 Times	118.38				
	More Than 15 Times	88.00				
3. Improvement	None	93.07	1.435	0.827	Fail to Reject Ho	Not Significant
	1-5 Times	92.82				
	6-10 Times	88.15				
	11-15 Times	99.59				
	More Than 15 Times	91.66				
4. Irrelevance	None	91.13	2.871	0.584	Fail to Reject Ho	Not Significant
	1-5 Times	90.78				
	6-10 Times	94.73				
	11-15 Times	107.00				
	More Than 15 Times	105.70				
5. Standards-Based	None	95.88	1.092	0.895	Fail to Reject Ho	Not Significant
	1-5 Times	91.56				
	6-10 Times	86.75				
	11-15 Times	108.88				
	More Than 15 Times	90.77				
6. Concept Development	None	94.79	1.483	0.830	Fail to Reject Ho	Not Significant
	1-5 Times	93.78				
	6-10 Times	86.42				
	11-15 Times	99.00				
	More Than 15 Times	78.32				
7. Formative Assessment	None	95.11	0.664	0.956	Fail to Reject Ho	Not Significant
	1-5 Times	92.40				
	6-10 Times	90.19				
	11-15 Times	100.25				
	More Than 15 Times	83.50				
8. Summative Assessment	None	90.95	0.994	0.911	Fail to Reject Ho	Not Significant
	1-5 Times	94.17				
	6-10 Times	88.61				
	11-15 Times	105.63				
	More Than 15 Times	84.64				

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 25 below displays that there is no significant difference between the respondents' conception of assessment and the number of trainings they attended. This is so because the p-values of all the items are greater than the 0.05 level of significance. The item formative assessment received the highest p-value of 0.978 which is greater than the 0.05 level of significance. Hence, null hypothesis was accepted. With the p-value of 0.196 which is also greater than the 0.05 level of significance, the item school accountability was decided to be of no significant difference in relation to respondents' trainings attended related to assessment.

This result implies that respondents' attendance to training related to assessment does not affect their conception of assessment.

5. Significant difference on the respondents in terms of assessment practices when they are grouped by profile

Table 26: Summary Table Showing the Significant Difference on the Respondent's Assessment Practices When Grouped According to Sex

Assessment Practices (Lacia, 2019)	Sex	Mean Rank	Mann-Whitney U	p-value	Decision	Remark
1. Purpose of Assessment	Male	84.21	2968.50	0.121	Fail to Reject Ho	Not Significant
	Female	95.68				
2. Classroom Assessment	Male	81.80	2846.00	0.032	Reject Ho	Significant
	Female	96.60				
3. Assessment Strategies	Male	84.53	2985.00	0.153	Fail to Reject Ho	Not Significant

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 26 shows the Mann-Whitney U test to confirm if significant difference between the respondents' assessment practices and sex. It shows that in classroom assessment there is a significant difference with a p-value of which is less than the 0.05 level of significance. On the other hand, there is no significant difference about purpose of assessment and assessment strategies as the two factors received p-values of 0.121 and 0.153, respectively.

This implies that sex influences the conduct of classroom assessment.

Table 27: Summary Table Showing the Significant Difference on the Respondent's Assessment Practices When Grouped According to Number of years in teaching

Assessment Practices (Lacia, 2019)	Number of years in teaching	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. Purpose of Assessment	1-3 Years	98.81	4.096	0.664	Fail to Reject Ho	Not Significant
	4-6 Years	89.92				
	7-9 Years	79.58				
	10-12 Years	89.10				
	13-15 Years	86.50				
	16-18 Years	93.83				
	19 Years & Above	95.97				
2. Classroom Assessment	1-3 Years	97.92	8.595	0.198	Fail to Reject Ho	Not Significant
	4-6 Years	94.73				
	7-9 Years	76.92				
	10-12 Years	106.70				
	13-15 Years	87.26				
	16-18 Years	89.00				
	19 Years & Above	78.25				
3. Assessment Strategies	1-3 Years	102.06	11.835	0.066	Fail to Reject Ho	Not Significant
	4-6 Years	92.28				
	7-9 Years	77.42				
	10-12 Years	102.00				
	13-15 Years	80.59				
	16-18 Years	40.00				
	19 Years & Above	83.13				

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Moreover, with the p-values that are greater than the 0.05 level of significance, result from Table 27 from the previous pages shows that there is no significant difference between the respondents' practices of assessment and the number of years they had been teaching.

This implies that the number of years in teaching of the respondents does not make any difference when doing or administering assessment.

Table 28: Summary Table Showing the Significant Difference on the Respondent's Assessment Practices When Grouped According to Subject Specialization

Assessment Practices (Lacia, 2019)	Subject specialization	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. Purpose of Assessment	AP	89.19	1.288	0.989	Fail to Reject Ho	Not Significant
	English	96.21				
	Filipino	92.10				
	MAPEH	91.53				
	Math	95.65				
	Science	84.67				
	TLE	93.63				
	EsP	90.20				
2. Classroom Assessment	AP	88.76	6.526	0.480	Fail to Reject Ho	Not Significant
	English	102.47				
	Filipino	90.38				
	MAPEH	86.80				
	Math	94.28				
	Science	87.65				
	TLE	80.31				
	EsP	109.20				
Assessment Practices (Lacia, 2019)	Subject specialization	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
3. Assessment Strategies	AP	83.24	4.204	0.756	Fail to Reject Ho	Not Significant
	English	97.24				
	Filipino	97.33				
	MAPEH	88.73				
	Math	100.30				
	Science	80.77				
	TLE	95.44				
	EsP	90.15				

Note*:* If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 28 from the previous page displays the result when Kruskal-Wallis H was utilized to verify if there exists a significant difference on the respondents' assessment practices when they were grouped according to subject specialization. "Purpose of Assessment, Classroom Assessment" and "Assessment Strategies" obtained p-values of 0.989, 0.480, and 0.756, respectively. Overall, the null hypothesis is accepted since all p-values are greater than 0.05 level of significance and therefore there is no significant difference between the variables.

Table 29: Summary Table Showing the Significant Difference on the Respondent's Assessment Practices When Grouped According to Highest Educational Attainment

Assessment Practices (Lacia, 2019)	Highest educational attainment	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. Purpose of Assessment	Master's (Grad.)	71.72	4.641	0.098	Fail to Reject Ho	Not Significant
	Master's (With Units Earned)	78.63				
	Bachelor's	94.39				
2. Classroom Assessment	Master's (Grad.)	103.72	1.472	0.479	Fail to Reject Ho	Not Significant
	Master's (With Units Earned)	84.85				
	Bachelor's	91.31				
3. Assessment Strategies	Master's (Grad.)	75.11	3.288	0.193	Fail to Reject Ho	Not Significant
	Master's (With Units Earned)	79.63				
	Bachelor's	94.00				

Note*:* If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Table 29 on the next page presents a no significant difference between the respondents' highest educational attainment and their assessment practices. The three items got the p-values of 0.098, 0.479 and 0.193 which are all greater than the 0.05 level of significance and thus, null hypotheses are accepted.

The result implies that the educational attainment of the respondents does not affect their actual practice of the assessment purpose, conduct of classroom assessment and actual choice and practice of assessment strategies.

Table 30: Summary Table Showing the Significant Difference on the Respondent's Assessment Practices When Grouped According to Position

Assessment Practices	Position	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. Purpose of Assessment	Classroom Teacher	96.73	10.234	0.006	Reject Ho	Significant
	Coordinator	82.43				
	Principal	48.36				
2. Classroom Assessment	Classroom Teacher	93.75	3.967	0.138	Fail to Reject Ho	Not Significant
	Coordinator	89.27				
	Principal	62.79				
3. Assessment Strategies	Classroom Teacher	97.20	10.849	0.005	Reject Ho	Significant
	Coordinator	81.22				
	Principal	47.07				

Note: If the p-value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

Shown in Table 30 on the succeeding page is the Kruskal-Wallis H test used to verify if there exists a significant difference on the respondents' assessment practices when they were grouped according to their position. The data showed that there is no significant difference about their assessment practices in terms of classroom assessment.

This implies that regardless of their position the practice of classroom assessment did not have a difference or position does not affect their conduct of classroom assessment. Meanwhile, it is worth noting that for items "purpose of assessment" and "assessment strategies", a significant difference emerged when they were compared to respondents' position. This means that as the respondents occupy different roles, there is a likelihood or possibility that the purpose of their assessment and at the same time their strategies may vary from one position to another.

Table 31: Summary Table Showing the Significant Difference on the Respondent's Assessment Practices When Grouped According to Number of trainings attended related to Assessment

Assessment Practices (Lacia, 2019)	Number of trainings attended related to Assessment	Mean Rank	Kruskal-Wallis H	p-value	Decision	Remark
1. Purpose of Assessment	None	89.23	0.644	0.958	Fail to Reject Ho	Not Significant
	1-5 Times	93.55				
	6-10 Times	89.89				
	11-15 Times	101.63				
	More Than 15 Times	97.14				
2. Classroom Assessment	None	87.89	2.146	0.709	Fail to Reject Ho	Not Significant
	1-5 Times	93.32				
	6-10 Times	98.83				
	11-15 Times	74.25				
	More Than 15 Times	100.18				
3. Assessment Strategies	None	89.26	4.252	0.373	Fail to Reject Ho	Not Significant
	1-5 Times	95.51				
	6-10 Times	100.28				
	11-15 Times	62.50				
	More Than 15 Times	75.45				

Note: If p value is less than or equal to the level of significance which is 0.05 reject the null hypothesis otherwise accept.

As shown on Table 31, respondents' attendance to training related to assessment has no significant difference on their assessment practices with the p-values that are greater than the 0.05 level of significance. This implies that the null hypothesis is accepted.

The above finding shows that respondents' assessment practices are not affected by the number of training related to assessment that they have attended.

B. Discussion

Below are important discussions from the data gathered in this study.

a) Out of the 184 teacher-respondents, there were 133 or 72.3% female while there were 51 or 27.7% of them were male. There were 77 or 41.8% of them teaching for 1-3 years, 32 or 17.4% of them are teaching for 4-6 years, 24 or 13% for 7-9 years, 17 or 9.2% for 13-15 years, 16 or 8.7% for 19 years and above, 15 or 8.2% for 10-12 year and 3 or 1.6% for 16-18 years. There were 34 or 18.5% of them with specialization in English, 29 or 15.8% in Araling Panlipunan, 27 or 14.7% in Math, 24 or 13% in Science, 24 or 13% in TLE, 20 or 10.9% in Filipino, 15 or 8.2% in MAPEH, 10 or 5.4% in Edukasyon sa Pagpapakatao and 1 or 0.5% has specialization in other subject. There were 146 or 79.3% are Bachelor's degree holders, 26 or 14.1% had earned units in Master's degree, 9 or 4.9% of them are Master's degree holders, 2, or 1.1% acquired Doctoral degree and 1 or 0.5% had earned units in Doctoral degree. There were 132 or 71.7% of them are classroom teachers, 43 or 23.4% are coordinators, 7 or 3.8% are principals and 2 or 1.1% occupy other positions. There were 105 or 57.1% of them who were able to attend 1-5 times, 46 or 25% of them have not attended training related

to assessment yet, 18 or 9.8% attended 6-10 times, 11 or 6% were able to attend training about assessment for more than 15 times and 4 or 2.2% attended 11-15 times.

b) Teachers' conceptions of assessment in terms of school accountability, the highest mean of 3.58 which is verbally interpreted as strongly agree was revealed to be the statements "Assessment provides information on how well schools are doing." and "Assessment shows the value schools add to student learning." while the lowest mean 3.48 verbally interpreted as strongly agree was "Assessment keeps schools honest and up-to-scratch."

In terms of student accountability, assessment conceived to determine if students meet qualifications standards got the highest mean of 3.53 which is verbally interpreted as strongly agree while "Assessment selects students for future education or employment." was the lowest mean of 3.32 and was still interpreted as "strongly agree".

Under the improvement factor, the highest mean was 3.56 was obtained by the item improvement-describe. On the other hand, improvement-validity with the mean of 3.19 was the lowest and verbally interpreted as agree.

Meanwhile, in terms of the factor irrelevance, it was revealed that the highest mean of 2.79 verbally interpreted as agree was the factor irrelevance accurate. The belief that assessment is important yet ignored got the lowest mean of 2.03 verbally interpreted as disagree.

As regard to the conception of assessment in terms of standards-based, the statement "Assessment determines learners' knowledge, skills and attitudes they need to demonstrate in every lesson and/or learning activity" obtained the highest mean of 3.55, verbally interpreted as strongly agree while the statement "Assessment answers the question, "What should the learners know?" got the lowest mean of 3.47 which was also verbally interpreted as strongly agree.

In terms of concept development, the item "The learner can generate meaning from oral, written and graphic messages" got the highest mean of 3.42 which was verbally interpreted as strongly agree. The statement "The learner can recall information and retrieve relevant knowledge from long-term memory" obtained the lowest mean which is 3.32 but still verbally interpreted as strongly agree.

Meanwhile, in terms of formative assessment, the highest weighted mean of 3.45 verbally interpreted as strongly agree, revealed that formative assessment helps prepare learners for summative assessments while formative assessment as informal got the lowest weighted mean of 2.95 verbally interpreted as agree.

In terms of summative assessment, summative assessment as an assessment of learning given at the end of a particular unit received the highest weighted mean of 3.61 verbally interpreted as strongly agree. With the lowest weighted mean of 2.96 verbally interpreted as agree was the conception that summative assessment is usually for the benefit of people other than the learner.

c) Regarding the level of assessment practices in terms of purpose of assessment, the statements "I ask questions or tasks that allow me to know whether students: can recall or remember what is taught in class" and "…explain Ideas and concepts" both received the highest weighted mean of 3.66 which were verbally interpreted as highly practiced and "I rank the students based on their class performance to inform other school officials" got the lowest weighted mean of 3.23 verbally interpreted as practiced.

In terms of classroom assessment, with a weighted mean of 3.72, verbally interpreted as highly practiced, the statement "I see to it that cheating is not encouraged in the classroom" got the highest weighted mean while the statement "I post the names of students who performed well in a test" got the lowest weighted mean of 3.10, verbally interpreted as practiced.

With regard to assessment strategies, Multiple Choice with 3.58 weighted mean was the highest which was verbally interpreted as highly practiced. On the contrary, Interviews and conferences" obtained the lowest mean of 3.07, verbally interpreted as practiced.

d) Regarding the difference on the conceptions of assessment when respondents were grouped according to sex, number of years in teaching, subject specialization, and number of trainings related to assessment attended, the study found out that there is no significant difference.

However, the difference on the conceptions of assessment in terms of school accountability (p-value= 0.004), student accountability (p-value= 0.036), improvement-student learning (p-value= 0.040) and formative assessment (p-value= 0.011) when respondents were grouped according to highest educational attainment was found significant since the p-values were all less than the 0.05 level of significance.

Additionally, a significant difference was also observed on the respondents' conceptions of assessment in terms of school accountability and standards-based when they were grouped according to position with the p-values of 0.036 and 0.044, respectively, all were less than the 0.05 level of significance.

e) Meanwhile, difference on the respondents' assessment practices in terms of classroom assessment when they were grouped according to sex found to be significant having a p-value of 0.032 which was less than the 0.05 level of significance. But, in terms of purpose of assessment and assessment strategies, there is no significant difference was found since the p-values 0.121 and 0.153, respectively, were greater than the 0.05 level of significance.

When grouped according to number of years in teaching, subject specialization, highest educational attainment, and number of trainings about assessment attended, no significant difference was observed between the respondents' assessment practices and the said profiles.

However, when they were grouped based on position a significant difference was found in terms of purpose of assessment and assessment strategies since the p-values were 0.006 and 0.005, respectively, which were all less than the 0.05 level of significance. Between classroom assessment and position was found to be of no significant difference since the p-value of 0.138 was greater than the 0.05 level of significance.

5. Conclusion

The researcher concludes the following:

a) Majority of the respondents are female, teaching for 4-6 years, specialized in English, bachelor's degree holders, classroom teachers and able to attend 1-5 trainings related to assessment.

b) Respondents strongly agreed on the conceptions of assessment in terms of school accountability, student accountability, improvement, standards-based, concept development, formative assessment and summative assessment. Disagreement was noted on irrelevance. This shows that teachers possess positive conceptions of

assessment. They did not see assessment as irrelevant or a bad practice.

c) The purpose of assessment and classroom assessment are highly practiced assessment practices than assessment strategies. This implies that teachers are more familiar with the purpose of assessment and administration of the classroom assessment than employing specific assessment strategies.

d) There was no significant difference between the conceptions of assessment when respondents were grouped according to sex, number of years in teaching, subject specialization, and number of trainings related to assessment attended. There was a significant difference on the conceptions of assessment in terms of school accountability, student accountability, improvement and formative assessment when respondents were grouped according to highest educational attainment. There was significant difference observed on the respondents' conceptions of assessment in terms of school accountability and standards-based when they were grouped according to position. This proves that upgrade in the teachers' educational attainment and promotion in their positions can help improve their conceptions of assessment.

e) There was a significant difference on the respondents' assessment practices in terms of classroom assessment when they were grouped according to sex. But, in terms of purpose of assessment and assessment strategies, there was no significant difference found. There was no significant difference found between the respondents' assessment practices and the number of years in teaching, subject specialization, highest educational attainment and number of trainings about assessment attended. There was significant difference found in terms of purpose of assessment and assessment strategies when the respondents were grouped according to position. But no significant difference in terms of classroom assessment when they were grouped based on position. This means that more diverse faculty members can provide more balanced assessment practices. Additionally, change in their positions can also help improve assessment practices.

6. Thank-you note

The author would like to thank previous researchers for their contributions in conducting research on the assessment conceptions and practices. Thanks to Mr. Gavin T. Brown, Dr. Joana D. Rural and Dr. Michelle Lacia, for their previous works were a big contribution to this study. The research results obtained by previous researchers, I and other readers can obtain complete information on the conceptions of assessment and practices as well. This information will certainly be very useful as a basis for further research to create innovative and useful resource for teachers. I hope that the results of writing this article can provide a broader picture of assessment and how teachers can maximize its purpose and use.

7. References

1. Arikunto S. Research Procedures: A Practical Approach. Indonesia: Rineka Cipta; 2006.
2. Looney A, Cumming J, van Der Kleij F, Harris K. Reconceptualizing the role of teachers as assessors: teacher assessment identity. Assessment in Education: Principles, Policy & Practice. 2018;25(5):442-467. DOI: 10.1080/0969594X.2016.1268090.
3. Atjonen P, Pöntinen S, Kontkanen S, Ruotsalainen P. Enhancing preservice teachers' assessment literacy: focus on knowledge base, conceptions of assessment, and teacher learning. Frontiers in Education. 2022;7:891391. DOI: 10.3389/feduc.2022.891391.
4. Azis A. Teachers' conceptions and use of assessment in student learning. Indonesian Journal of Applied Linguistics. 2012;2(1):40–51. DOI: 10.17509/ijal.v2i1.72.
5. Azis A. Conceptions and practices of assessment: a case of teachers representing improvement conception. TEFLIN Journal. 2015;26(2):129-154. DOI: 10.15639/teflinjournal.v26i2/129-154.
6. Avvisati F. How are school performance and school climate related to teachers' experience? PISA in Focus, No. 88. Paris: OECD Publishing; 2018. DOI: 10.1787/af992283-en.
7. Baird M. Exploring Teachers' Perceptions of Assessment Literacy in a High Performing School District in New Jersey [dissertation]. South Orange (NJ): Seton Hall University; 2021. Available from: https://scholarship.shu.edu/cgi/viewcontent.cgi?article=4032&context=dissertations
8. Ballada CJA, Alino MAC. Exploring Filipino teachers' conceptions of assessment. Educational Measurement and Evaluation Review. 2018 Jul:1-15.
9. Barnes N, Fives H, Dacey CM. U.S. teachers' conceptions of the purposes of assessment. Teaching and Teacher Education. 2017;65:107–116. DOI: 10.1016/j.tate.2017.02.017.
10. Bernardo AB. Formative assessment and school improvement: prospects and challenges in the development of formative assessment that could help improve student learning outcomes in Philippine schools. Educational Measurement and Evaluation Review. 2018:1-12.
11. Black P, William D. Classroom assessment and pedagogy. Assessment in Education: Principles, Policy & Practice. 2018;25(1):7-74. DOI: 10.1080/0969594X.2018.1441180.
12. Brown G. Teachers' conceptions of assessment: implications for policy and professional development. Assessment in Education: Principles, Policy & Practice. 2004;11(3):301-318. DOI: 10.1080/0969594042000304609.
13. Caldone D. The Tensions of Teacher Assessment Identity: What Secondary Teachers Think, Do, and Learn from Experience [dissertation]. United States: ProQuest Dissertations Publishing; 2021. Available from: https://search.proquest.com/openview/44f16a512c98f5f14aec2ae0c166e48e/1?pq-origsite=gscholar&cbl=18750&diss=y
14. Celebre-Baird M. Exploring Teachers' Perceptions of Assessment Literacy in a High Performing School District in New Jersey [dissertation]. South Orange (NJ): Seton Hall University; 2021. Available from: https://scholarship.shu.edu/dissertations/2930
15. Choong C. Definition of Assessment, Its Purpose in Classroom Assessment and Assessment Techniques [Internet]. 2013 [cited 2024]. Available from: https://www.academia.edu/4935982/Definition_of_Ass

essment_Its_Purpose_in_Classroom_Assessment_and_Assessment_Techniques
16. CNN Philippines. PH Grade 4 students worst in math and science proficiency: study [Internet]. 2020 Dec 10 [cited 2024]. Available from: https://www.cnnphilippines.com/news/2020/12/10/PH-Grade-4-students-worst-in-math-and-science-proficiency--study-.html
17. Dayal HC, Lingam GI. Fijian teachers' conceptions of assessment. Australian Journal of Teacher Education. 2015;40(8):42-58. Available from: https://files.eric.ed.gov/fulltext/EJ1073634.pdf
18. DeLuca C, Klinger D. Assessment literacy development: identifying gaps in teacher candidates' learning. Assessment in Education: Principles, Policy & Practice. 2010;17(4):419-438.
19. DeLuca C, LaPointe-McEwan D, Luhanga U. Teacher assessment literacy: a review of international standards and measures. Educational Assessment, Evaluation and Accountability. 2016;28(3):251–272.
20. Department of Education. Policy guidelines on classroom assessment for the K to 12 basic education program. DepEd Order No. 8, s. 2015. Philippines: Department of Education; 2015.
21. Department of Education. Policy guidelines on system assessment in the K to 12 basic education program. DepEd Order No. 29, s. 2017. Philippines: Department of Education; 2017.
22. Department of Education. National adoption and implementation of the Philippine professional standards for teachers. DepEd Order No. 42, s. 2017. Philippines: Department of Education; 2017.
23. Dixson DD, Worrell FC. Formative and summative assessment in the classroom. Theory into Practice. 2016;55(2):153–159. DOI: 10.1080/00405841.2016.1148989.
24. Doyle A, Conroy Johnson M, Donlon E, McDonald E, Sexton P. The role of the teacher as assessor: developing student teacher's assessment identity. Australian Journal of Teacher Education. 2021;46(12):1-18. DOI: 10.14221/ajte.2021v46n12.4.
25. Edutopia. Why is assessment important? [Internet]. 2008 [cited 2024]. Available from: https://www.edutopia.org/assessment-guide-importance
26. Huber SG, Skedsmo G. Assessment in education—from early childhood to higher education. Educational Assessment, Evaluation and Accountability. 2016;28(3):201–203. DOI: 10.1007/s11092-016-9245-x.
27. Lacia M. Classroom practices in mathematics: effects on elementary and secondary school students' achievement in mathematics in Region XII, Philippines. International Journal of Scientific and Technology Research. 2019;8(10):1-8.
28. Looney A, Cumming J, van Der Kleij F, Harris K. Reconceptualising the role of teachers as assessors: teacher assessment identity. Assessment in Education: Principles, Policy & Practice. 2017;25(5):442-467. DOI: 10.1080/0969594X.2016.1268090.
29. Magno C, Ocampo J. Can students value for formative assessment develop their academic self-regulation and grit? Educational Measurement and Evaluation Review. 2018:1-15.
30. Malaluan N. Learning assessment to play crucial role in post-COVID education [Internet]. BusinessWorld. 2022 Oct 9 [cited 2024]. Available from: https://www.bworldonline.com/opinion/2022/10/09/479353
31. Melaku T, Wudu M. Primary school teachers' conceptions and practices of assessment and their relationships. Cogent Education. 2022;9(1):2090185. DOI: 10.1080/2331186X.2022.2090185.
32. Mertler CA, Campbell C. Measuring Teachers' Knowledge & Application of Classroom Assessment Concepts: Development of the "Assessment Literacy Inventory" [Internet]. 2005 [cited 2024]. Available from: https://eric.ed.gov/?id=ED490355
33. Manitoba Education. Rethinking Classroom Assessment with Purpose in Mind. Winnipeg: Manitoba Education, Citizenship and Youth; 2006. Available from: https://www.edu.gov.mb.ca/k12/assess/wncp/full_doc.pdf
34. Mirian BM, Zulnaidi H. Mathematics teachers' conceptions of assessment: gender and academic qualification comparisons. International Journal of Instruction. 2020;13(2):239–252. DOI: 10.29333/iji.2020.13217a.
35. Monteiro V, Mata L, Santos NN. Assessment conceptions and practices: perspectives of primary school teachers and students. Frontiers in Education. 2021;6:631185. DOI: 10.3389/feduc.2021.631185.
36. OECD. Evaluation and assessment frameworks for improving school outcomes. Paris: OECD Publishing; 2009. Available from: https://www.oecd.org/education/school/46927511.pdf
37. Opre D. Teachers' conceptions of assessment. Procedia - Social and Behavioral Sciences. 2015;209:229–233. DOI: 10.1016/j.sbspro.2015.11.222.
38. Ounis T. Exploring secondary teachers' perceptions of classroom assessment in a Tunisian context. International Journal of Language and Literature. 2017;4(2):148-158. Available from: http://ijllnet.com/journals/Vol_4_No_2_June_2017/15.pdf
39. Piosang TL. A cross-sectional analysis of classroom assessment literacy of English teachers in secondary and tertiary levels. Educational Measurement and Evaluation Review. 2017;8(1):1-15.
40. Queen's University. Assessment [Internet]. Kingston (ON): Queen's University Centre for Teaching and Learning; [cited 2024]. Available from: https://www.queensu.ca/teachingandlearning/modules/assessments/index.html
41. Randel B, Apthorp H, Beesley AD, Clark TF, Wang X. Impacts of professional development in classroom assessment on teacher and student outcomes. The Journal of Educational Research. 2016;109(5):491-502.
42. Remesal A. Primary and secondary teachers' conceptions of assessment: a qualitative study. Teaching and Teacher Education. 2011;27(2):472-482.
43. Rural J. Teachers' conceptions of assessment using Brown's four-factor model and DepEd assessment policy: implications for policy review. Turkish Journal of Computer and Mathematics Education. 2019;10(3):765-785. Available from: https://turcomat.org/index.php/turkbilmat/article/download/2170/1890/4087
44. Siarova H, Sternadel D, Masidlauskaite R. Assessment

practices for 21st century learning: review of evidence. NESET II report. Luxembourg: Publications Office of the European Union; 2017. DOI: 10.2766/71491.
45. Schleicher A. PISA 2018 Insights and Interpretations. Paris: OECD Publishing; 2019. Available from: https://www.oecd.org/pisa/publications/pisa-2018-results.htm
46. Siddiqi GS. Assessment practices in education [Internet]. 2014 [cited 2024]. Available from: https://www.academia.edu/23813478/Assessment_Practices_in_Education
47. Smith L, Hill M, Cowie B, Gilmore A. Preparing teachers to use the enabling power of assessment. In: Wyatt-Smith C, Klenowski V, Colbert P, editors. Designing Assessment for Quality Learning. Dordrecht: Springer; 2014. p. 303-323. DOI: 10.1007/978-94-007-5902-2_19.
48. Sulaymonova NJ. Learning and assessment. JournalNX - A Multidisciplinary Peer Reviewed Journal. 2021;7(7):412–414. Available from: https://repo.journalnx.com/index.php/nx/article/view/2950
49. Tesorio PJC, Canizares MJF. Philippine K-12 secondary science teachers assessment conceptions in promoting student learning. The Normal Lights. 2018;12(2):1-32. Available from: https://po.pnuresearchportal.org/ejournal/index.php/normallights/article/view/1002
50. Top Hat. Authentic Assessment [Internet]. Toronto (ON): Top Hat; [cited 2024]. Available from: https://tophat.com/glossary/a/authentic-assessment/
51. Unal A, Unal Z. An examination of K-12 teachers' assessment beliefs and practices in relation to years of teaching experience. Georgia Educational Researcher. 2019;16(1):1-25.
52. UNESCO. Assessment [Internet]. Paris: UNESCO; [cited 2024]. Available from: UNESCO website.
53. Vraketta G. Classroom assessment practices and the enhancement of student learning [Internet]. 2013 [cited 2024]. Available from: https://www.academia.edu/4496178/Classroom_Assessment_Practices_and_the_Enhancement_of_Student_Learning
54. Yamtim V, Wongwanich S. A study on classroom assessment literacy of primary school teachers. Procedia - Social and Behavioral Sciences. 2014;116:2998-3004.
55. Yazon A. Statistics: A Guide for Researchers and Thesis Writers. Quezon City: Wiseman's Books Trading, Inc.; 2020.