Randwick International of Education and Linguistics Science (RIELS) Journal
Vol. 6, No. 2, June 2025 | Page 553-573 |
https://www.randwickresearch.com/index.php/rielsj

Classroom-Based Assessment Practices and Challenges of Primary Teachers: Basis for A Proposed Upskilling Program

DOI: https://doi.org/10.47175/rielsj.v6i2.1175

| Maricel M. Bravo[1,*] | Rosemarie D. Paceño[2] |

[1,2] Agusan Colleges Incorporated, Butuan City, Philippines

[*] maricel.bravo11@gmail.com

ABSTRACT

The study explored the classroom-based assessment (CBA) practices and challenges of primary teachers in the West Butuan District as the basis for a proposed upskilling program. Employing a mixed-methods approach, the investigation utilized a validated survey questionnaire administered to 44 primary teachers, supplemented by a focus group discussion (FGD) to capture contextual and experiential insights. Quantitative data were analyzed using weighted mean and Spearman Rho correlation, while qualitative responses were thematically coded. Findings indicated that while primary teachers demonstrated a generally high level of knowledge in key CBA domains—Table of Specifications (TOS), item development, test administration, scoring, and item evaluation—gaps remained in the practical application of this knowledge, particularly in ensuring test coverage, managing assessment time, and conducting item analysis. Statistically significant correlations were found between knowledge in all five CBA domains and actual assessment practices. The FGD further revealed persistent challenges, including limited training in item evaluation, difficulty in constructing time-bound assessments, and misalignment between test items and instructional objectives. In response, a three-day face-to-face upskilling program was developed to address these gaps through collaborative workshops, real-time feedback, and hands-on output creation, with the goal of improving assessment quality, enhancing instructional planning, and promoting better learner outcomes.

Keywords

classroom-based assessment; assessment practices; primary education, item analysis; TOS; teacher professional development; upskilling program

INTRODUCTION

Teachers play a vital role in shaping learners' learning, with support from administrators, parents, and the wider school community. Primary teachers carry the main responsibility for helping students build strong foundations in knowledge and skills (Allen, 2019). To do this effectively, teachers regularly assess learners' progress through test results, classroom participation, and observed behaviors. Creative and engaging teaching strategies are especially important in encouraging active learning and motivation among young learners.

The role of primary teachers is even more critical in addressing learners' diverse learning needs and in promoting values like critical thinking (Department of Education, 2020). However, while Classroom-based Assessments (CBA) are essential for tracking learning, some teachers need further support to use these tools effectively. This is because some

teachers may lack updated training in designing assessments aligned with current curriculum standards, struggle to interpret assessment data for instructional decision-making, or face time constraints that limit the ability to develop quality assessment tools.

DeLuca et al. (2021) emphasize that many teachers feel underprepared in assessment literacy, particularly in creating formative assessments and using results to guide instruction. This highlights the need for continuous professional development to enhance instructional decisions and learner support. Similar challenges are evident in classrooms, where accurately designing assessments to reflect learning progress remains a concern. Smith and Holloway (2022) add that limited professional development often leads teachers to rely on traditional testing methods, restricting the use of innovative, learner-centered assessment practices. Without institutional support and access to modern tools, teachers may struggle to move beyond outdated approaches, affecting student engagement and growth. This highlights the need for sustained capacity-building efforts that empower teachers to adopt more relevant and responsive assessment strategies.

This study addressed the low results of quarterly assessment scores of primary learners at Gabriel Elementary School in the West District, Division of Butuan City during 2023, revealing a noticeable discrepancy between current actual performance of learners, as measured by the quarterly assessment scores, and the academic standards set by the curriculum. The issue appeared to be that existing assessment practices were not effectively supporting improvements in learner scores, which may have contributed to these low results. Given that similar challenges were likely present in other schools within the district, this study examined classroom assessment practices across nine schools to identify specific areas for improvement. Effective assessment practices are found to be crucial, as they not only helped identify student struggles but also informed teaching strategies and drove academic improvements. The study emphasized the need for an upskilling program to equip teachers with the necessary tools and strategies to better engage students and improve the academic performance.

Theoretical Framework

This study was anchored on Albert Bandura's Self-Efficacy Theory (Bandura, A. 1977) which emphasizes that individuals are more likely to engage in behaviors perceived as achievable. Within the educational context, teacher self-efficacy refers to the perceived capability to carry out instructional tasks effectively, including classroom-based assessment (CBA). Teachers who believe in the competence to design, administer, and evaluate assessments are more likely to apply effective assessment strategies in actual classroom practice. This theory supports the investigation of teachers' knowledge in areas such as Table of Specifications, item development, test administration, scoring, and item evaluation, as well as how this knowledge translates into assessment practices. Self-efficacy also provides a lens for understanding the challenges teachers encounter—particularly how low confidence in specific assessment tasks may hinder performance and impact student learning outcomes.

Furthermore, Thomas R. Guskey's Theory of Teacher Professional Development (Guskey, 2002) emphasizes that effective professional development must be structured, intentional, and sustained to bring about meaningful improvements in teaching and learning. This theory underpins the proposed upskilling program, which is specifically designed to effectively address the specific assessment challenges identified in the study. By focusing on practical, collaborative training, the program aims to strengthen teachers' knowledge and skills, foster sustained professional growth, and enhance the overall quality of classroom-based assessment practices.

Objective of the Study

This study explored the classroom-based assessment practices and challenges of primary teachers in the West District, Division of Butuan City.

RESEARCH METHODS

Research Design

The study utilized a mixed methods design, combining quantitative and qualitative approaches. The quantitative component employed a descriptive-correlational design, using cross-sectional data gathered through a survey instrument to examine the relationship between teachers' knowledge and practices in classroom-based assessment. Complementing this, the qualitative component utilized a Focus Group Discussion (FGD) to gain deeper insights into teachers' experiences and the challenges they face in implementing assessment practices.

The descriptive aspect provided an overview of teachers' current assessment practices, knowledge levels, and the challenges they encounter in the classroom. Meanwhile, the correlational analysis explored how these practices were related to their level of knowledge and the specific difficulties they faced. By integrating both quantitative and qualitative data, the mixed methods approach allowed for a comprehensive and nuanced understanding of the gaps in teachers' assessment practices, capturing not only measurable trends and relationships but also the underlying reasons, perceptions, and contextual factors that influence how assessments are conducted in real classroom settings.

Research Locale

The study was conducted at the West District, Division of Butuan City. The West Butuan District was composed of nine (9) elementary schools, namely, Gabriel Elementary School, Bliss Elementary School, Libertad Central Elementary School, Pareja Integrated School, Masao Elementary School, Lumbocan Elementary School, Ambago Central Elementary School, Pinamaculan Elementary School, and Dalingdingan Elementary School, which were known for their diverse population and educational challenges.

The schools represented a mix of urban and rural settings, allowing for a comprehensive understanding of the classroom-based assessment practices across different contexts. Given their proximity to Butuan City proper, the researcher used public transportation, such as tricycles or multi-cabs, to reach the study sites. The variation in transportation options also allowed for flexibility in accessing the schools based on weather conditions and availability.

Research Instrument

The study used a researcher-made survey questionnaire as the primary tool for data collection. The questionnaire was divided into two sections: the first section assessed the level of knowledge of primary teachers on classroom-based assessment, covering key areas such as the table of specifications, item development, test administration, scoring, and item evaluation. The second section explored how teachers manifested classroom-based assessment practices, specifically in terms of coverage, time management, and test design. This structured approach allowed for a clear and comprehensive analysis of teachers' knowledge and practices in classroom-based assessment.

In addition to the survey questionnaire, a Focus Group Discussion (FGD) was conducted to gather qualitative data. The FGD aimed to provide in-depth insights into the challenges and experiences of teachers regarding the implementation of classroom-based assessments. This mixed-methods approach, combining the quantitative survey with qualitative insights

from the FGD, enabled a more comprehensive understanding of the factors influencing teachers' assessment practices.

The survey questionnaire was designed to be straightforward and easy to understand, ensuring that all participants could provide accurate and consistent responses. The inclusion of the FGD further enriched the data, offering a deeper context to the survey findings.

Data Gathering Procedure

Following the validation of the instrument, data collection began in a structured manner. The validated survey, along with a request for approval, was submitted to the district supervisor to ensure alignment with guidelines. Once approved, the survey was distributed to school principals for coordination with primary teachers.

The survey was administered either online or in print, depending on school preferences. Clear instructions were provided to ensure accurate responses, with an emphasis on confidentiality to encourage honest participation.

In addition to the survey, a Focus Group Discussion (FGD) was conducted with selected teachers to gain deeper insights into their experiences and challenges with classroom-based assessments. To ensure confidentiality and maintain anonymity, participants in the FGD were assigned numerical identifiers (e.g., P1, P2, P3, etc.), with the numbering system randomly assigned and not correlated with any specific participant characteristics. The FGD responses were recorded, transcribed, and analyzed qualitatively.

Data Analysis Plan

Statistical Treatment

This study used the following statistical treatment in analyzing the data. Frequencies and Percentages. These were used to describe the distribution of the population and respondent of the study.

Weighted Means. These were computed to determine the average responses on teachers' knowledge and assessment practices, summarizing how frequently or effectively specific practices were applied.

Spearman Rho Correlation Analysis. Spearman Rho was used to examine the relationship between teachers' knowledge of classroom-based assessment and the extent to which these practices were implemented. This test helped determine whether a significant association existed between the two variables.

RESULTS AND DISCUSSION

Table 1. Level of Knowledge of the Primary Teachers on Classroom-based Assessment in terms of constructing the Table of Specifications

Indicators	Wtd Mean	SD	Verbal Description	Knowledge Level
1. The Table of Specifications (TOS) ensures alignment of classroom assessments with learning objectives and curriculum standards.	4.84	0.370	Strongly Agree	Very Satisfactory
2. The TOS is used to organize assessment items across different topics and cognitive levels in a structured and balanced way.	4.73	0.451	Strongly Agree	Very Satisfactory

3. A well-developed TOS includes a fair mix of lower-order thinking skills (LOTS) (e.g., remembering, understanding, applying) and higher-order thinking skills (HOTS) (e.g., analyzing, evaluating, creating).	4.64	0.487	Strongly Agree	Very Satisfactory
4. Assessments developed using the TOS effectively promote critical thinking, problem-solving, and analytical reasoning.	4.57	0.625	Strongly Agree	Very Satisfactory
5. Regular review and updates of the TOS are conducted based on student performance data to improve assessment quality.	4.50	0.550	Strongly Agree	Very Satisfactory
Overall Weighted Mean	**4.65**	**0.389**	**Strongly Agree**	**Very Satisfactory**

Legend: 1.00-1.49-Strongly Disagree/Very Poor; 1.50-2.49-Disagree/Poor; 2.50-3.49-Uncertain/Fair; 3.50 4.49-Agree/Satisfactory; 4.50-5.00-Strongly Agree/Very Satisfactory

Table 1 shows the highest-rated indicator, with a weighted mean of 4.84, is "The Table of Specifications (TOS) ensures alignment of classroom assessments with learning objectives and curriculum standards." This suggests that teachers have a strong understanding of how the TOS serves to align assessments with curriculum standards. Teachers confidently apply the TOS to ensure that assessments are consistent with learning outcomes. In the FGD, teachers shared that they follow a systematic process to ensure alignment. As P3 explained, "We always start with the MELCs and make sure each topic has fair coverage," while P8 noted, "We use a school-provided TOS template to guide us." This reinforces the teachers' consistent approach to ensuring that assessments align with the curriculum.

The lowest, though still very satisfactory, mean rating of 4.50 was recorded for the indicator on regularly reviewing and updating the TOS based on learner performance data. This indicates that while the teachers recognize the importance of this practice, it may not yet be consistently applied in all actual classroom settings. The slightly higher standard deviation (0.550) suggests that responses varied more for this indicator, highlighting the challenges teachers face in regularly updating the TOS. During the FGD, teachers indicated that revisions to the TOS are infrequent, often prompted by poor student performance or feedback from school heads. As P5 remarked, "Only when the test results are low or the coordinator notices a misalignment, we go back to revise the TOS." This insight points to a gap between understanding the importance of regular updates and the actual practice of doing so, mainly due to time constraints and competing priorities.

The overall weighted mean of 4.65 (SD = 0.389) suggests that, overall, teachers have a very satisfactory level of knowledge in constructing the TOS. The low standard deviation indicates consistency in the responses, which suggests a shared understanding and application of the TOS framework. However, the challenge lies in the consistent and timely revision of the TOS. This is further emphasized in the FGD, where teachers expressed a need for ongoing support to improve TOS preparation. P9 noted, "We need more concrete examples and collaborative sessions to review each other's TOS," highlighting a desire for professional development and resources to refine the practices.

Despite a strong theoretical understanding, teachers face difficulties when trying to apply the TOS in real-time, particularly when it comes to revising it regularly based on student data. Teachers cited the vagueness of some learning competencies and the difficulty of matching them to appropriate cognitive levels in Bloom's taxonomy as a challenge. P1 mentioned, "Some competencies are too broad—it's hard to decide where they fit in Bloom's taxonomy," while P6 pointed out that not all learning objectives are easily assessed through

written tests. These challenges further underline the complexities teachers encounter when constructing and revising the TOS, and the need for clearer guidelines and more practical resources to support their work.

These findings are consistent with existing literature, such as the work of Santos (2023), which discusses the importance of continuous reflection and adaptation in assessment practices. Despite the teachers' solid foundation in constructing the TOS, the ability to revise and adapt it regularly remains hindered by workload pressures and a lack of structured systems for periodic updates. This highlights the need for more targeted support and professional development to help teachers bridge the gap between knowledge and practice, especially in aligning assessments with evolving student needs and curriculum changes.

Table 2 presents the responses of the primary teachers from the West Butuan District regarding the knowledge of classroom-based assessment, specifically focusing on item development.

Table 2. Level of Knowledge of the Primary Teachers on Classroom-based Assessment in terms of Item Development

Indicators	Wtd Mean	SD	Verbal Description	Knowledge Level
1. Test items are created to align with the learning objectives and assess a range of cognitive levels, from recall to higher-order thinking.	4.84	0.370	Strongly Agree	Very Satisfactory
2. Test items are clear, unbiased, and free from unnecessary complexity to accurately measure students' understanding.	4.55	0.548	Strongly Agree	Very Satisfactory
3. A variety of question types are designed to assess different skills and learning outcomes.	4.59	0.583	Strongly Agree	Very Satisfactory
4. Test items are written to be appropriate for the students' grade level and designed to avoid cultural or linguistic bias.	4.61	0.579	Strongly Agree	Very Satisfactory
5. Test items are specific and relevant to the content taught, providing a fair and accurate assessment of student learning.	4.61	0.493	Strongly Agree	Very Satisfactory
Overall Weighted Mean	**4.64**	**0.392**	**Strongly Agree**	**Very Satisfactory**

Legend: 1.00-1.49-Strongly Disagree/Very Poor; 1.50-2.49-Disagree/Poor; 2.50-3.49-Uncertain/Fair; 3.50-4.49-Agree/Satisfactory; 4.50-5.00-Strongly Agree/Very Satisfactory

The data presented in Table 2 shows the highest-rated indicator, with a weighted mean of 4.84, is "Test items are created to align with the learning objectives and assess a range of cognitive levels, from recall to higher-order thinking." This reflects a solid understanding among teachers of the need to design assessments that are not only aligned with learning objectives but also challenge students to demonstrate a broad range of cognitive abilities. Teachers recognize the importance of creating assessments that foster critical thinking and require students to apply knowledge in more complex ways. This emphasis on higher-order thinking (HOTS) was further emphasized by P4 during the FGD, who said, "It's easier to write recall-type questions. Creating analysis or evaluation questions takes time and creativity." This highlights the intentionality behind their efforts to challenge students beyond simple recall.

The lowest-rated indicator, though still very satisfactory, is "Test items are clear, unbiased, and free from unnecessary complexity to accurately measure students' understanding," with a weighted mean of 4.55. The slightly higher standard deviation (0.548) suggests variability in responses, indicating that while teachers understand the need for

clarity and fairness in test items, the practical challenges of achieving this are recognized. Time constraints and workload pressures were recurring themes in the FGD, with P7 noting, "We are overloaded with tasks, so we end up reusing or slightly editing previous test items," illustrating how workload affects the consistency and clarity of newly developed items. This supports Caluya's (2020) observation that many teachers rely on textbook-based or previously used test items, often at the expense of promoting higher-order thinking skills. Teachers acknowledged that while they strive for clear and unbiased items, external factors often hinder their ability to consistently create high-quality test questions.

The overall weighted mean of 4.64 (SD=0.392) reflects a very satisfactory level of knowledge among teachers in item development. The relatively low standard deviations across the indicators suggest that the teachers have a shared understanding of the principles of effective item development. This consistency is reflected in the findings from FGD participants, who indicated a collaborative approach to developing test items. As P2 shared, "I write the test items alone first, then we compare them with grade-level partners," which shows that teachers rely on collaboration for feedback, even though this is often done after the initial draft.

While the teachers in West Butuan District demonstrate strong knowledge in developing test items, challenges related to time and resources were frequently cited in the FGD. For example, P6 pointed out, "Sometimes we help each other, but everyone is too busy with their own workload." Teachers expressed a need for more dedicated time for collaboration, which would further improve the quality of test items. This is consistent with the literature on assessment practices, such as Caluya (2020), Baysa (2021), and Santos (2023), who note that despite improvements in test item development, constraints like time pressures, lack of resources, and the challenge of creating higher-order thinking items remain ongoing concerns.

In conclusion, the results from Table 2 align with the findings from the FGD, indicating that while teachers are well-versed in the fundamentals of item development, they face significant challenges in ensuring clarity, fairness, and the creation of HOTS items due to time and resource constraints. These insights provide valuable context for understanding the challenges teachers face in classroom assessment practices, which will be addressed in the proposed upskilling program.

Table 3. Level of Knowledge of the Primary Teachers on Classroom-based Assessment in terms of Test Administration

Indicators	Wtd Mean	SD	Verbal Description	Knowledge Level
1. Standardized procedures are followed for administering assessments, including preparing materials, providing instructions, and monitoring the testing environment.	4.57	0.546	Strongly Agree	Very Satisfactory
2. Test instructions are explained clearly to students before they begin the assessment.	4.55	0.589	Strongly Agree	Very Satisfactory
3. Accommodations are provided for students with special needs during test administration to ensure accessibility.	4.36	0.685	Agree	Satisfactory
4. A positive and conducive environment is created for administering classroom-based assessments, ensuring that students are comfortable and focused.	4.52	0.505	Strongly Agree	Very Satisfactory

5. Classroom management strategies are implemented to minimize distractions and maintain a secure testing environment during assessments.	4.52	0.505	Strongly Agree	Very Satisfactory
Overall Weighted Mean	**4.50**	**0.437**	**Strongly Agree**	**Very Satisfactory**

Legend: 1.00-1.49-Strongly Disagree/Very Poor; 1.50-2.49-Disagree/Poor; 2.50-3.49-Uncertain/Fair; 3.50-4.49-Agree/Satisfactory; 4.50-5.00-Strongly Agree/Very Satisfactory

The data presented in Table 3 indicate that the highest weighted mean of 4.57 with a standard deviation of 0.546 was obtained from the indicator, "Standardized procedures are followed for administering assessments, including preparing materials, providing instructions, and monitoring the testing environment." This is interpreted as "Strongly Agree" with a Very Satisfactory level of knowledge. This result suggests that the respondents demonstrate a high level of familiarity and adherence to standardized assessment protocols during test administration. In actual school settings, this is reflected during formal assessments, where teachers ensure the readiness of materials, deliver clear and consistent instructions, and maintain a controlled and focused environment for learners. The finding affirms existing classroom practices, which may be attributed to the continuous implementation of DepEd guidelines and policies.

On the other hand, the indicator with the lowest weighted mean is "Accommodations are provided for students with special needs during test administration to ensure accessibility," with a mean of 4.36 and a standard deviation of 0.685, interpreted as "Agree" with a Satisfactory level of knowledge. Although still positive, this suggests a slightly lower confidence or experience in applying inclusive practices during assessments. From observations in classrooms, while many teachers express willingness to accommodate learners with special needs, they sometimes lack the specific knowledge or strategies to do so effectively. This may be attributed to limited exposure to specialized training in inclusive education or insufficient resources in regular classrooms.

Overall, the domain of test administration earned an Overall Weighted Mean of 4.50 with a standard deviation of 0.437, interpreted as "Strongly Agree" and rated as Very Satisfactory. This indicates that, in general, teachers demonstrate a strong understanding of how to properly administer classroom-based assessments. Their ability to maintain order, follow assessment protocols, and create a conducive environment supports the reliability and fairness of student evaluations.

These findings support the discussions presented in the reviewed literature. Magsino (2019) identified logistical issues in test administration, including classroom overcrowding and a lack of materials, which can hinder the proper implementation of standardized procedures. However, the very satisfactory rating observed suggests that despite these challenges, teachers are making deliberate efforts to uphold standardized assessment practices.

Table 4. Level of Knowledge of the Primary Teachers on Classroom-based Assessment in terms of Scoring

Indicators	Wtd Mean	SD	Verbal Description	Knowledge Level
1. Sufficient guidance is provided on selecting and applying appropriate scoring methods (e.g., rubrics, grading scales) for classroom-based assessments.	4.55	0.589	Strongly Agree	Very Satisfactory

2. There is level of confidence in grading both objective and subjective items (e.g., multiple-choice, essays, short answers) fairly.	4.36	0.574	Agree	Satisfactory
3. Clear scoring rubrics or criteria are designed for test items to ensure grading consistency and objectivity.	4.39	0.689	Agree	Satisfactory
4. Assessment scores and feedback on strengths and areas for improvement are communicated to students and parents in a clear and understandable manner.	4.50	0.550	Strongly Agree	Very Satisfactory
5. Training on grading assessments for fairness and reliability has been received and applied to extent.	4.45	0.663	Agree	Satisfactory
Overall Weighted Mean	**4.45**	**0.490**	**Agree**	**Satisfactory**

Legend: 1.00-1.49-Strongly Disagree/Very Poor; 1.50-2.49-Disagree/Poor; 2.50-3.49-Uncertain/Fair; 3.50-4.49-Agree/Satisfactory; 4.50-5.00-Strongly Agree/Very Satisfactory

The data presented in Table 4 reveals that the highest weighted mean of 4.55 with a standard deviation of 0.589 was obtained from the indicator, "Sufficient guidance is provided on selecting and applying appropriate scoring methods (e.g., rubrics, grading scales) for classroom-based assessments." This was interpreted as "Strongly Agree" with a Very Satisfactory level of knowledge. This finding indicates that the respondents are highly knowledgeable about choosing and applying scoring techniques suitable for different types of assessments. In actual classroom practice, this is evident as many primary teachers have developed the habit of using rubrics, especially for performance tasks and written outputs, to ensure objective and standardized scoring. This finding affirms real classroom conditions, likely due to past trainings, Learning Action Cell (LAC) sessions, or school-based initiatives that introduced teachers to different scoring tools and methods. The high rating suggests that teachers understand the importance of matching assessment tools with the nature of student tasks, which contributes to transparency and fairness in grading.

On the other hand, the lowest weighted mean was recorded in the item, "There is level of confidence in grading both objective and subjective items (e.g., multiple-choice, essays, short answers) fairly," with a mean of 4.36 and a standard deviation of 0.574, interpreted as "Agree" and rated as Satisfactory. While the result still reflects a positive response, it indicates slightly lower confidence among teachers when it comes to scoring different types of assessment items, particularly those that require subjective judgment. From observations, teachers tend to be more at ease with objective tests like multiple-choice questions, while they sometimes express hesitation in scoring essays or open-ended responses due to fear of bias or uncertainty in applying scoring criteria consistently. This may be attributed to a lack of practice or limited exposure to strategies in constructing and scoring subjective tests.

The overall weighted mean for the scoring domain is 4.45 with a standard deviation of 0.490, interpreted as "Agree" with a Satisfactory level of knowledge. This suggests that, in general, the 44 primary teachers from the West Butuan District possess a solid foundational understanding of scoring practices in classroom-based assessment. However, while the knowledge is acceptable and consistent across most indicators, the slight variations in the data point to specific areas needing improvement, particularly in subjective grading and consistent application of scoring tools. This is a realistic reflection of current classroom practices where, despite strong intentions and understanding, teachers may struggle to apply scoring techniques across diverse assessment formats.

The results presented in Table 4 reinforce existing literature on the importance of structured and consistent scoring practices in classroom-based assessments. Dela Peña

(2020) highlighted that scoring often lacks consistency when rubrics are not utilized, a challenge that some teachers may still encounter, particularly when evaluating subjective tasks. Salandanan (2021) emphasized that rubrics play a critical role in ensuring fairness, especially for performance-based assessments. However, in schools like Gabriel Elementary School, manual scoring and subjective judgment remain prevalent, which may contribute to variations in grading. In classroom experiences, it has been observed that the lack of consistent use of rubrics often leads to discrepancies in scoring, with teachers relying on their own interpretations. These insights confirm that while teachers possess a foundational understanding of scoring principles, there is a clear need to strengthen the capacity in developing and using rubrics effectively.

Table 5. Level of Knowledge of the Primary Teachers on Classroom-based Assessment in terms of Item Evaluation

Indicators	Wtd Mean	SD	Verbal Description	Knowledge Level
1. Assessment items are regularly evaluated to ensure they effectively measure the specific learning objectives they are intended to assess.	4.57	0.587	Strongly Agree	Very Satisfactory
2. Efforts are made to identify potential biases in assessment items and ensure fairness and equity for all students.	4.61	0.538	Strongly Agree	Very Satisfactory
3. Assessment items are reviewed and revised based on feedback and student performance data to improve their effectiveness.	4.50	0.629	Strongly Agree	Very Satisfactory
4. Assessment items are examined to ensure they challenge students appropriately at different cognitive levels and provide meaningful measures of learning progress.	4.43	0.587	Agree	Satisfactory
5. Item analysis is conducted to evaluate the difficulty of test items, identifying those that are too easy or too difficult, and adjustments are made to improve the effectiveness.	4.39	0.722	Agree	Satisfactory
Overall Weighted Mean	**4.50**	**0.500**	**Strongly Agree**	**Very Satisfactory**

Legend: 1.00-1.49-Strongly Disagree/Very Poor; 1.50-2.49-Disagree/Poor; 2.50-3.49-Uncertain/Fair; 3.50-4.49-Agree/Satisfactory; 4.50-5.00-Strongly Agree/Very Satisfactory

Table 5 presents the highest mean score, 4.61, was recorded under the indicator, “Efforts are made to identify potential biases in assessment items and ensure fairness and equity for all learners.” This implies a strong awareness among teachers of the need for fairness in test construction. During the FGD, participants emphasized the intent to avoid content that could disadvantage certain groups of learners. Teachers noted that they make a conscious effort to use inclusive language and culturally sensitive content. This strong result aligns with the Department of Education’s current emphasis on gender responsiveness, inclusive education, and equitable assessment practices.

In contrast, the lowest mean score was 4.39 with a relatively high standard deviation of 0.722 for the indicator, “Item analysis is conducted to evaluate the difficulty of test items, identifying those that are too easy or too difficult, and adjustments are made to improve the effectiveness.” This was interpreted as “Agree” with a Satisfactory knowledge level. The wider spread of responses suggests that this is a less consistently practiced aspect of assessment among teachers. This gap was echoed in the FGD, where teachers admitted they rarely conduct formal item analysis. P1 shared, “If too many students fail a question, I know

there's a problem with it," indicating reliance on intuition over systematic evaluation. P8 further admitted, "I rarely check item difficulty unless required." These statements point to a clear gap in the technical understanding and application of item evaluation procedures.

The overall weighted mean for this domain is 4.50, interpreted as "Strongly Agree" and reflective of a Very Satisfactory level of knowledge. This suggests that teachers recognize the importance of item evaluation in theory, particularly in terms of aligning test items with objectives and promoting fairness. However, the FGD responses indicate that actual practices are limited by lack of training and time constraints. As P3 noted, "We don't know how to compute difficulty and discrimination indices." Furthermore, tools and statistical procedures related to item analysis are rarely emphasized in school-based training. Teachers typically revise items only after poor student performance reveals flaws. P5 stated, "I change questions only if many got them wrong and I realize the question wasn't clear," confirming that revisions are reactionary rather than proactive.

Participants also expressed a desire for hands-on training on item analysis, preferably using digital tools such as Excel. P9 remarked, "Hands-on training in analyzing test data will really help us improve." This suggests that while the knowledge base for fairness and alignment is present, teachers are seeking more practical, skills-based support to deepen the capacity for evaluating and improving test quality.

These findings are consistent with Ramirez and Banzon (2019), who found that item evaluation is often neglected in actual classroom settings due to time limitations and lack of technical know-how. Likewise, Tolentino (2022) emphasized that even basic item analysis can significantly enhance assessment reliability and validity. The present study confirms that while West Butuan District teachers show strong intent and awareness, the actual conduct of item evaluation—particularly in terms of item analysis—is underdeveloped. In practice, test items are often reused without revision, and evaluation is limited to surface-level reflections on student performance. This highlights the need for capacity-building interventions that focus on equipping teachers with simple, user-friendly tools and techniques for analyzing test data and improving assessment validity.

Table 6. Summary Table: Level of Knowledge of Primary Teachers on Classroom-Based Assessment

Domain	Overall Weighted Mean	Standard Deviation (SD)	Verbal Description	Knowledge Level
Table of Specifications (TOS)	4.65	0.389	Strongly Agree	Very Satisfactory
Item Development	4.64	0.392	Strongly Agree	Very Satisfactory
Test Administration	4.50	0.437	Strongly Agree	Very Satisfactory
Scoring	4.45	0.490	Agree	Satisfactory
Item Analysis & Evaluation	4.47	0.482	Agree	Satisfactory
Overall	**4.54**	**0.478**	**Strongly Agree**	**Very Satisfactory**

Table 6 summarizes the primary teachers' level of knowledge across five domains of classroom-based assessment. The highest knowledge level was observed in the Table of Specifications (M=4.65; SD=0.389) and Item Development (M=4.64; SD=0.392), both rated as Very Satisfactory (Strongly Agree). These results suggest that teachers demonstrate strong knowledge in these areas. On the other hand, the lowest knowledge level was observed in Scoring (M=4.45; SD=0.490) and Item Analysis & Evaluation (M=4.47; SD= 0.482), which were rated as Satisfactory (Agree). Overall, the primary teachers' knowledge

across all domains was rated as Very Satisfactory, with an overall weighted mean of 4.54, indicating a strong proficiency in classroom-based assessment practices, though some areas may benefit from targeted improvement efforts. This highlights the need for focused professional development in specific assessment domains to further enhance teachers' competencies.

Extent of Manifestation of Classroom-based Assessment Practices

This section focuses on how teachers apply their assessment knowledge in actual classroom practices. It presents findings on the extent to which classroom-based assessment is implemented, specifically in the areas of coverage, time management, and test design. Each area is discussed highlighting both the strengths and areas that may need improvement.

Table 7. Extent of manifestation of Classroom-based Assessment Practices in terms of Coverage

Indicators	Wtd Mean	SD	Verbal Description	Manifestation Level
1. The required content is consistently covered within the available assessment time.	4.52	0.505	Always	Very High
2. Classroom assessments are effectively aligned with the curriculum standards.	4.61	0.538	Always	Very High
3. A balance is maintained between covering all necessary content and assessing key learning objectives within the given assessment time frame.	4.39	0.655	Almost Always	High
4. Learners are adequately prepared for assessments covering a broad range of material.	4.23	0.605	Almost Always	High
5. Assessments are adapted as needed to remain aligned with the required curriculum standards, even when changes to the curriculum occur.	4.52	0.549	Always	Very High
Overall Weighted Mean	**4.45**	**0.419**	**Almost Always**	**High**

Legend: 1.00-1.49-Never/Very Low; 1.50-2.49-Rarely/Low; 2.50-3.49-Sometimes/Moderately HIgh; 3.50-4.49-Almost Always/ High; 4.50-5.00-Always/Very High

Table 7 in the previous page presents the extent of manifestation of classroom-based assessment practices in terms of Coverage. The highest mean among the indicators for coverage was recorded in the statement "Classroom assessments are effectively aligned with the curriculum standards," which obtained a weighted mean of 4.61 and a standard deviation of 0.538. This was verbally described as “Always” and interpreted as a Very High level of manifestation. This result reflects a strong commitment among primary teachers to ensure that classroom assessments are directly based on curriculum standards. This finding aligns with actual teaching and assessment practices observed in the West District. Schools consistently emphasize curriculum alignment during lesson planning and assessment preparation. Teachers frequently refer to the Most Essential Learning Competencies (MELCs) to ensure that assessments measure what has been taught. This consistency may be attributed to regular instructional monitoring, strong school leadership, and structured planning routines that reinforce curriculum-based assessment practices. The finding is affirmative of real circumstances, as curriculum alignment has been a longstanding instructional priority across the division.

The lowest mean was found in the indicator "Learners are adequately prepared for assessments covering a broad range of material," which had a weighted mean of 4.23 and a standard deviation of 0.605, verbally described as “Almost Always” and interpreted as a high level of manifestation. Although this indicates generally effective preparation, it suggests that not all learners are consistently ready for assessments that span a wide range

of topics. This result may be linked to several factors. One possible cause is the varying learning pace among students, which can make it challenging to cover broad content areas effectively within limited instructional time. Additionally, sudden interruptions in the school calendar, such as suspensions or unplanned activities, can affect lesson pacing and the depth of preparation. In some cases, learners with limited foundational skills may require more intensive review, making comprehensive coverage more difficult. The finding reflects a real but manageable challenge in classroom settings, highlighting the need for more targeted review strategies and differentiated instruction before assessment periods.

Overall, the results yielded an overall weighted mean of 4.45 with a standard deviation of 0.419, verbally described as "Almost Always" and interpreted as a High level of manifestation. This indicates that classroom-based assessment practices in terms of coverage are generally strong and consistently implemented. This high level of manifestation can be attributed to the deliberate efforts of teachers to align assessments with curriculum requirements, ensure content coverage within time constraints, and adapt assessments as necessary. The ability to adjust assessment content in response to curriculum changes also shows flexibility and commitment to maintaining relevance. Despite minor challenges in student preparedness, the findings affirm that teachers maintain high standards in planning and delivering assessments that are comprehensive, curriculum-based, and focused on essential learning outcomes.

The findings in Table 7 support and expand on earlier insights from previous studies. Javier (2019) pointed out that assessments often focus on simpler aspects of the curriculum, which may lead to the neglect of broader learning goals. While the current study shows a high level of alignment between assessments and curriculum standards, the slightly lower mean on learner preparedness suggests that, in some cases, content breadth may still be a challenge. Villanueva (2021) emphasized the importance of aligning assessments with the Most Essential Learning Competencies (MELCs), a practice that is evidently being upheld in the West District. However, in some schools, assessments tend to prioritize content that is easier to deliver or evaluate, which risks excluding more complex but essential competencies.

Table 8. Extent of manifestation of Classroom-based Assessment Practices in terms of time Management

Indicators	Wtd Mean	SD	Verbal Description	Manifestation Level
1. The amount of time available for preparing assessments has allowed for the development of thorough and complex assessment items.	4.50	0.591	Always	Very High
2. Enough time in test item development have helped to focus assessments on the most essential learning competencies.	4.39	0.538	Almost Always	High
3. Balancing instructional time with assessment preparation has allowed for sufficient time for both tasks.	4.27	0.694	Almost Always	High
4. The time allocated for administering assessments has been adequate to fully assess students' understanding of the material.	4.25	0.719	Almost Always	High
5. Time spent on grading assessments has provided the opportunity to offer comprehensive feedback to students.	4.27	0.694	Almost Always	High
Overall Weighted Mean	**4.34**	**0.499**	**Almost Always**	**High**

Legend: 1.00-1.49-Never/Very Low; 1.50-2.49-Rarely/Low; 2.50-3.49-Sometimes/Moderately HIgh; 3.50-4.49-Almost Always/ High; 4.50-5.00-Always/Very High

Table 8 presents the extent of manifestation of classroom-based assessment practices in terms of Time Management. Among the indicators of time management, the highest mean was recorded in the statement "The amount of time available for preparing assessments has allowed for the development of thorough and complex assessment items," with a weighted mean of 4.50 and a standard deviation of 0.591, verbally described as “Always” and interpreted as a Very High level of manifestation. This indicates that primary teachers consistently allocate sufficient time for developing high-quality assessments. This finding aligns with common classroom practices in the West District, Division of Butuan City, where many teachers are observed to prioritize the planning and preparation of assessment tools. During scheduled planning sessions, teachers often engage in collaborative development of assessment items, ensuring alignment with the Most Essential Learning Competencies (MELCs). This consistency may be attributed to strong administrative support, structured planning routines, and ongoing professional development activities such as Learning Action Cell (LAC) sessions. The finding reflects real circumstances in the field, where assessment quality is increasingly emphasized as a key component of effective instruction.

On the other hand, the lowest mean was observed in the indicator "The time allocated for administering assessments has been adequate to fully assess students’ understanding of the material," which received a weighted mean of 4.25 and a standard deviation of 0.719, verbally described as “Almost Always” and interpreted as a High level of manifestation. This suggests that while assessment administration is generally well-managed, certain constraints affect its consistency. Observations across primary classrooms show that the limited time for each subject, coupled with overlapping school activities and varying student pace, can pose challenges in fully administering assessments. Learners who require more time, particularly in subjects such as reading or mathematics, may not always be accommodated within the standard class schedule. This condition likely contributes to the slightly lower rating of this indicator. The finding is reflective of actual classroom conditions and underscores the need for more flexible scheduling and time management approaches during assessment periods.

Overall, the results yielded an overall weighted mean of 4.34 with a standard deviation of 0.499, verbally described as “Almost Always” and interpreted as a High level of manifestation. This indicates that, in general, teachers in the West District effectively manage their time in conducting classroom-based assessments. The high level of manifestation may be attributed to sustained efforts in professional development and a culture of instructional planning that emphasizes both teaching and assessment. Despite the heavy workload and additional responsibilities placed on teachers, assessment-related tasks remain a consistent priority. This finding affirms that the current time management practices are functioning effectively, although there remains a need to enhance the administration phase to ensure all learners are given equal opportunities to demonstrate their understanding.

The findings in Table 8 align with and reinforce existing insights from prior studies. Alvarado (2020) emphasized that limited planning time often results in lower-quality assessments, a concern that remains relevant given the varying levels of time management practices observed. Although teachers in the West District demonstrated a high level of manifestation in balancing instructional and assessment preparation time, the slightly lower mean on time allocated for administering assessments and providing feedback indicates ongoing challenges. Rosales and Co (2022) highlighted the importance of incorporating time management strategies into teacher development programs, which is particularly significant in settings where teaching loads and administrative tasks limit planning time. In many schools, such constraints have been observed to impact assessment quality.

Table 9. Extent of manifestation of Classroom-based Assessment Practices in terms of Test Design

Indicators	Wtd Mean	SD	Verbal Description	Manifestation Level
1. The alignment of assessments with lesson objectives ensures the effective coverage of essential content.	4.70	0.509	Always	Very High
2. A broad range of cognitive skills, from basic recall to critical thinking, is measured through the assessment items.	4.52	0.590	Always	Very High
3. Clear and measurable evidence of student learning outcomes is demonstrated through the assessments.	4.41	0.622	Almost Always	High
4. A balanced mix of question types is created to address diverse learning needs.	4.52	0.628	Always	Very High
5. Training in assessment design enhances the quality and effectiveness of classroom assessments.	4.39	0.689	Almost Always	High
Overall Weighted Mean	**4.51**	**0.517**	**Always**	**Very High**

Legend: 1.00-1.49-Never/Very Low; 1.50-2.49-Rarely/Low; 2.50-3.49-Sometimes/Moderately HIgh; 3.50-4.49-Almost Always/ High; 4.50-5.00-Always/Very High

Table 9 presents the extent of manifestation of classroom-based assessment practices among primary teachers in terms of Test Design. Based on the data, the highest mean was observed in the indicator "The alignment of assessments with lesson objectives ensures the effective coverage of essential content," which obtained a weighted mean of 4.70 and a standard deviation of 0.509. This was verbally described as “Always” and interpreted as a Very High level of manifestation. This result indicates that teachers consistently align assessment items with lesson objectives, ensuring that essential content is adequately assessed. This finding is strongly affirmed by current classroom practices in the West District, where curriculum alignment is a central focus of lesson planning and assessment development. Teachers are regularly guided to align assessments with the Most Essential Learning Competencies (MELCs) during Learning Action Cell (LAC) sessions and through instructional supervision. The emphasis on alignment ensures that assessments are purposeful and directly reflect what has been taught, resulting in more valid measures of student learning. The consistency of this practice may be driven by both policy direction and internal school monitoring systems that check for coherence between instruction and assessment.

Conversely, the lowest mean was seen in the indicator "Training in assessment design enhances the quality and effectiveness of classroom assessments," which recorded a weighted mean of 4.39 and a standard deviation of 0.689, verbally described as “Almost Always” and interpreted as a High level of manifestation. While still rated highly, this suggests that access to or application of training in test design is not fully consistent among teachers. This may be due to varying opportunities for teachers to participate in formal training sessions specifically focused on assessment design. Although some training is provided through division-led or school-based programs, the depth and frequency of such sessions may not be sufficient to address all teachers’ needs. Additionally, without continuous follow-through or mentoring, the application of advanced assessment design strategies may vary. This finding may not fully reflect the ideal, suggesting a gap between the perceived benefits of training and the actual availability or impact of such interventions.

Overall, the data showed a weighted mean of 4.51 with a standard deviation of 0.517, verbally described as “Always” and interpreted as a Very High level of manifestation. This

indicates that test design practices among primary teachers in the West District are consistently strong and well-implemented. Emphasis on a range of cognitive skills, a balanced mix of question types, and clarity of learning outcomes shows a matured understanding of assessment design principles. The finding is consistent with observed instructional practices, where teachers regularly craft assessment tools that challenge learners' higher-order thinking skills. Nonetheless, the result highlights the importance of sustaining and expanding capacity-building programs to further enhance assessment literacy among educators.

The findings in Table 9 align with prior studies. Briones (2021) highlighted that poorly designed tests can confuse learners and fail to assess key skills. This is addressed in the study by the strong alignment of assessments with lesson objectives and the use of diverse question types. However, the slightly lower mean regarding training in test design indicates room for improvement in teacher training. Santos (2023) noted that workshops can improve test clarity, diversity, and fairness, directly influencing student performance. Many teachers still face challenges in creating fair and effective assessments, which can impact learner outcomes.

Table 10. Summary Table: Extent of Manifestation of Classroom-Based Assessment Practices

Domain	Overall Weighted Mean	Standard Deviation (SD)	Verbal Description	Manifestation Level
Coverage	4.45	0.419	Almost Always	High
Time Management	4.34	0.499	Almost Always	High
Test Design	4.51	0.517	Always	Very High
Overall	**4.4**	**0.478**	**Almost Always**	**High**

Table 10 indicates that teachers generally demonstrate a high level of consistency in their assessment practices, with test design showing the highest manifestation at a Very High level (M=4.51). However, coverage and time management were rated as High (M=4.45 and 4.34, respectively), suggesting that while teachers are performing well, there are areas that could benefit from further refinement. These results emphasize the need for targeted professional development to equip teachers with the skills to manage classroom-based assessments effectively.

Test of Significant Relationship between Teachers' Level of Knowledge and their Classroom-based Assessment Practices

Table 4 on the next page shows the Spearman Rho correlation between teachers' knowledge in assessment domains and the practices in coverage, time management, and test design. Results indicate significant positive relationships, suggesting that greater knowledge leads to consistent assessment practices.

Table 11. Spearman Rho Correlation analysis between level of Knowledge and classroom-based assessment practices

		Coverage	Time Management	Test Design
Table of Specifications	Correlation Coefficient	.435**	.457**	.473**
	p-value	.003	.002	.001
	Decision on H_o	Reject H_o	Reject H_o	Reject H_o
	Interpretation	Significant	Significant	Significant

Item Development	Correlation Coefficient	.684**	.650**	.671**
	p-value	.000	.000	.000
	Decision on H_o	Reject H_o	Reject H_o	Reject H_o
	Interpretation	Significant	Significant	Significant
Test Administration	Correlation Coefficient	.713**	.707**	.640**
	p-value	.000	.000	.000
	Decision on H_o	Reject H_o	Reject H_o	Reject H_o
	Interpretation	Significant	Significant	Significant
Scoring	Correlation Coefficient	.709**	.643**	.643**
	p-value	.000	.000	.000
	Decision on H_o	Reject H_o	Reject H_o	Reject H_o
	Interpretation	Significant	Significant	Significant
Item Evaluation	Correlation Coefficient	.732**	.713**	.598**
	p-value	.000	.000	.000
	Decision on H_o	Reject H_o	Reject H_o	Reject H_o
	Interpretation	Significant	Significant	Significant

***significant @ p <.01*

Table 11 shows the result of the correlation analysis between the level of the teachers' knowledge on classroom-based assessment and the extent of manifestation of the assessment practices. It can be gleaned from the data in the Table that the level of knowledge of the teachers in terms of the construction of the Table of Specifications (TOS) has a significant relationship with the assessment practices along coverage(ρ=.435; p=.003); along time management (ρ=.457; p=.002); and along test design (ρ=.473; p=.001). This finding indicates that those teachers who have higher levels of knowledge on table of specifications have the tendencies to have very good practices in classroom assessment as they are able to ensure that the test items are consistent with the lesson coverage and the competencies that are to be assessed. In terms of time management, teachers with strong TOS knowledge effectively allocate time for developing and administering assessments. This ensures that the necessary amount of time is available for preparing thorough assessment items and balancing instructional time, as highlighted in the time management practices. Teachers also ensure that adequate time is spent on grading, allowing them to provide valuable feedback to learners. Regarding test design, a well-constructed TOS enables teachers to design assessments that align with lesson objectives, cover essential content, and measure a range of cognitive skills. This approach ensures that assessments are not only varied but also effectively evaluate student understanding, as reflected in the test design practices.

The results of this study align with literature highlighting the importance of teachers' knowledge in constructing the table of specifications (Tan & Lim, 2019; Gempes, 2020). Teachers with strong TOS knowledge showed better time management and test design practices, supporting Briones' (2021) call for continuous training. Studies by Alvarado (2020) and Rosales and Co (2022) also emphasized time management in assessment planning, which this study confirms. These findings reinforce the need for practical, localized upskilling programs (Cayanan & Gamboa, 2021) to improve classroom-based assessment.

The level of knowledge of the teachers in terms of item development has shown significant relationship with the practices along coverage (ρ=.684; p=.000); time management (ρ=.650; p=.000) and test design (ρ=.671; p=.000).The way teachers design the

test items are generally perceived to be aligned with the curriculum standards, hence, the most essential learning targets are the very competencies that are covered in the assessments. Moreover, item development has also to consider the time needed by the learners to answer each item, so, there is the practice to graduate the items from simple to more challenging items. It has also been observed that, while there is a suggested test design, some teachers would vary test formats to meet the target of assessing the learners' competencies along the different cognitive levels of learning. Additionally, teachers adjust item difficulty based on previous experiences and aim to include questions that assess both recall and application of concepts. They incorporate problem-solving items to encourage critical thinking, ensuring assessments are balanced and aligned with the curriculum's cognitive demands.

This supports the findings of Tan and Lim (2019) and Gempes (2020), who emphasized the importance of training in item development. Similarly, Mendoza (2022) and Lim and Luna (2021) found that knowledgeable teachers apply better assessment practices, which aligns with the observed classroom implementation in this study.

Knowledge of classroom assessment in terms of test administration also shows significant relationship with assessment practices along coverage (ρ=.713; p=.000); time management (ρ=.707; p=.000); and test design (ρ=.640; p=.000). The teachers are already used to testing the most essential learning competencies for a particular assessment period. This knowledge allows them to consider better test administration strategies that will ensure that all the preparations done in the initial stages of assessment are administered in the most advantageous manner for all the learners. Furthermore, teachers apply strategies such as clear test instructions, organized seating arrangements, and time reminders during test administration to help maximize focus and reduce anxiety among learners. They also ensure that the testing environment is conducive to concentration, minimizing distractions to uphold fairness. These practices support the smooth flow of assessment, promote efficient time use, and help learners perform at their best, aligning with the principles of effective coverage and test design.

These practices align with the findings of Magsino (2019), who identified logistical issues that hinder effective test administration, such as overcrowding and insufficient materials. Similarly, Alvarado (2020) highlighted how inconsistencies in test administration impact fairness, echoing the importance of standardized protocols in your findings. By understanding the relationship between teachers' knowledge of test administration and the practices, it becomes clear how professional development, as suggested by Briones (2021), is essential for maintaining consistency and fairness in assessments.

In terms of the teachers' knowledge in scoring, the Table 4 analysis shows relationship with coverage (ρ=.709; p=.000); time management (ρ=.643; p=.000); and test design (ρ=.643; p=.000). The scoring scheme of the teachers are based on some standard transmutation of scores and equivalent descriptive rating. Additionally, teachers ensure that their scoring practices are consistent and aligned with the learning objectives, using rubrics or scoring guides especially for constructed response items. This helps maintain objectivity and fairness in grading. Some teachers also allocate time for analyzing item results to identify patterns of student errors, which in turn informs future item revisions. These practices not only enhance the reliability of assessment results but also support efficient time use and reinforce alignment between test design and intended learning outcomes.

Teachers' proficiency in scoring methods like rubrics and grading scales aligns with findings of Dela Peña (2020) and Salandanan (2021), which highlighted the importance of rubrics in ensuring consistency and fairness in assessments. The significant relationship between scoring knowledge and effective practices mirrors Lim and Luna's (2021) finding

that strong assessment knowledge leads to better teaching outcomes. These insights underline the importance of an upskilling program to improve teachers' assessment practices.

Finally, Table 4 shows the result of the correlation analysis on knowledge of the teachers in classroom-based assessment in terms of item evaluation shows to have significant relationship with their assessment practices along coverage (ρ=.732; p=000); time management (ρ=713; p=.000); and test design(ρ=.598; p=.000). After the test, the teachers are required to determine the mastery levels of the learners to determine if there will be interventions or remediations that needed to be done. While teachers recognize the importance of evaluating assessment items for fairness and alignment, they often face time constraints and lack of training in conducting thorough evaluations.

The findings of the study, which show a significant relationship between teachers' knowledge in item evaluation and the assessment practices in terms of coverage, time management, and test design, align with the work of Tan and Lim (2019), who noted the lack of sufficient understanding in constructing and using assessment tools like the Table of Specification (TOS). Gempes (2020) also highlighted that teacher training in assessment practices, such as item evaluation, significantly improves the quality and reliability of assessments and suggests strengthening teachers' skills in evaluating test items.

Based on the Focus Group Discussion insights, a three-day face-to-face upskilling program on the next page titled "Enhancing Classroom-Based Assessment Practices and Addressing Challenges among Primary Teachers in the West Butuan District" is proposed to address the recurring issues identified. The program aims to bridge the gap between theoretical knowledge and practical application through collaborative, hands-on learning experiences.

CONCLUSION

Based on the findings of the study the following conclusions were drawn;

1. While primary teachers have strong foundational knowledge in classroom-based assessment, there remain conceptual and practical gaps, especially in applying item evaluation techniques and updating tools like the TOS.
2. Teachers generally implement classroom-based assessments effectively, but real-world constraints such as limited time, large class sizes, and system-imposed formats hinder optimal application.
3. The statistically significant relationship between knowledge and practice underscores the importance of continuous teacher training, as improved knowledge tends to result in more effective assessment strategies.
4. Persistent challenges such as difficulties in crafting higher-order questions, limited use of student performance data, and infrequent revision of assessment tools highlight the need for ongoing support, structured feedback mechanisms, and collaborative professional development.
5. A targeted upskilling program is essential to bridge the gap between knowledge and consistent application, ensuring sustainable improvements in classroom-based assessment. Such a program fosters continuous teacher development, ultimately enhancing instructional quality and student learning outcomes.

Recommendations

Based on the conclusions of the study, the following recommendations are offered for considerations;

1. To support school heads and teachers in reinforcing CBA knowledge, professional development programs should continue to focus on foundational areas such as the Table of Specifications, item development, test administration, scoring, and item evaluation. While knowledge levels are already strong, training should now emphasize deeper understanding and flexible application, especially in item evaluation where teachers admitted gaps. This would benefit both teachers' competence and learners' outcomes through better-aligned assessments.
2. To assist teachers in enhancing assessment implementation, practical workshops should be conducted on optimizing assessment coverage, managing time effectively, and designing tests that are both structured and adaptable. School administrators should provide support mechanisms—such as planning time and sample materials—to help teachers deal with large class sizes and limited preparation time. These improvements will create more consistent, inclusive, and efficient classroom assessments.
3. Given the significant link between knowledge and practice, school leaders should institutionalize ongoing mentoring and coaching programs. These initiatives should help teachers transfer their CBA knowledge into more consistent assessment behavior, especially under challenging work conditions. Doing so promotes long-term improvements in teaching strategies and supports learners' academic growth.
4. In response to the persistent challenges revealed through the FGD, particularly in TOS updating, collaborative item writing, and item analysis, it is recommended that school heads facilitate regular Professional Learning Communities (PLCs). These PLCs should serve as a venue for peer feedback, shared resource development, and reflective discussion. Teachers would benefit from reduced isolation in assessment planning, ultimately leading to more robust assessment tools and practices.
5. To ensure the sustainability and refinement of the upskilling program, an evaluation framework should be developed and implemented by school administrators in collaboration with future researchers. This should include teacher feedback, classroom observation, and student performance monitoring. Such a system would guide the continuous improvement of training efforts and serve as a reference point for future studies on teacher development and classroom assessment innovation.

REFERENCES

Alvarado, J. M. (2020). Managing assessment practices in large classrooms: Challenges and strategies. *Philippine Journal of Basic Education*, *12*(1), 45–53.

Allen, T. D. (2019). The role of educators in shaping 21st-century learners. Education and Development Review, 15(2), 34–47.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215.

Baysa, R. A. (2021). Enhancing item-writing skills of teachers through professional learning communities. *Journal of Educational Research and Training, 9(*1), 78–90.

Briones, M. L. (2021). The impact of test design on learner outcomes in elementary education. *International Journal of Assessment and Evaluation, 11*(3), 66–74.

Caluya, R. C. (2020). Textbook dependence and item development practices among elementary teachers. *Education Studies Review, 7*(2), 105–117.

Cayanan, A. R., & Gamboa, L. M. (2021). Designing relevant teacher training: A needs-based approach. *Philippine Teacher Education Journal, 5*(1), 89–102.

Dela Peña, J. F. (2020). Assessment scoring: Aligning rubrics with performance tasks in elementary schools. *Southeast Asian Journal of Pedagogical Studies, 8*(1), 23–35.

DeLuca, C., O'Connor, P., & Collet, K. (2021). Teacher assessment literacy and its impact on classroom practices.

Department of Education. (2020). DepEd Order No. 11, s. 2020: Adoption of the Basic Education Learning Continuity Plan for School Year 2020-2021 in light of the COVID-19 public health emergency. https://www.deped.gov.ph

Department of Education. (2020). DepEd Memorandum No. 162, s. 2020: Reaffirming the importance of formative assessment in the K to 12 Basic Education Program. https://www.deped.gov.ph

Gempes, G. P. (2020). Training needs assessment for enhancing classroom assessment competency. *International Journal of Research in Education and Psychology,* 5(3), 99–109.

Guskey, T. R. (2002). Professional development and teacher change. Teachers and Teaching: Theory and Practice, gus 381–391. https://doi.org/10.1080/135406002100000512

Javier, M. C. (2019). The practice of assessment coverage in the implementation of the K to 12 curriculum. *Education Quarterly, 4*(2), 134–147.

Konig, J., Jäger-Biela, D. J., & Glutsch, N. (2020). Adapting to online teaching during COVID-19 school closure: Teacher education and teacher competence effects among early career teachers in Germany. *European Journal of Teacher Education, 4*3(4), 608–622. https://doi.org/10.1080/02619768.2020.1809650.

Lim, C. R., & Luna, J. A. (2021). Exploring the link between assessment literacy and classroom practice. *Asian Journal of Educational Research, 9*(4), 208–220.

Magsino, L. M. (2019). Classroom-based testing issues in overcrowded schools. *Journal of Educational Issues in Developing Countries, 6*(2), 54–63.

Mendoza, P. R. (2022). Bridging the gap between assessment knowledge and practice: A case study of elementary teachers. *Philippine Journal of Education, 18*(1), 77–88.

Ramirez, H. T., & Banzon, M. E. (2019). Teachers' utilization of item analysis in test evaluation. *Journal of Assessment and Curriculum Development, 3*(2), 45–56.

Rosales, A. G., & Co, M. V. (2022). The effect of standardized test administration on learner achievement. *Southeast Asian Journal of Educational Practice, 7*(3), 120– 132.

Salandanan, G. G. (2021). Effective rubric design for elementary performance assessments. *Teacher Education Today, 5*(4), 91–104.

Santos, E. D. (2023). Localized upskilling workshops and their effect on assessment practices. *Journal of Teacher Development and Instruction, 11*(1), 55–70.

Smith, A., & Holloway, M. (2022). Challenges in implementing innovative assessment practices in primary education.

Tan, R. E., & Lim, M. D. (2019). Table of Specification: Its relevance in the K-12 assessment landscape. *Education and Measurement Journal, 10*(3), 88–99.

Tolentino, A. B. (2022). Enhancing assessment practices through post-test evaluation and scoring technology. *Asian Journal of Learning Assessment, 4*(2), 101–114.

Villanueva, P. A. (2021). Curriculum-aligned assessment practices among elementary teachers. *Philippine Journal of Curriculum Studies, 6*(1), 33–47.