Original Article

Integration of Authentic Assessment in Teaching Science and Academic Performance among Students

Joyna Jamero [1,*], Jenelyn Sabellina [2]

Received: 19 April 2026; Revised: 01 June 2026;
Accepted: 03 June 2026; Published: 06 June 2026

DOI: https://doi.org/10.66074/ITX22Q44W3

Abstract

Authentic assessment plays a crucial role in Science education by focusing on real-world applications, performance-based tasks, and higher-order thinking skills. This study investigated the extent of authentic assessment (AA) integration among Science teachers in the GUINSACAT District and its relationship with students' academic performance. Using a descriptive-correlational design, data were gathered from a total population of 40 Science teachers, and students' first-quarter grades for the 2025-2026 school year served as the metric for academic performance. Results showed that teachers demonstrated a "High Extent" of AA integration, frequently utilizing performance tasks, real-life applications, and student feedback. Students' performance was categorized as "Very Satisfactory" with an average grade of 85.33. However, the study revealed no significant relationship between the extent of authentic assessment integration and students' grades ($r_s = -0.006$, $p = 0.971$). Similarly, there were no significant differences in assessment practices across age, teaching experience, and educational attainment groups. While Science education often emphasizes the value of performance-based tasks, these findings indicate that within the current grading framework of this cohort, teacher-reported assessment integration does not statistically correlate with quarterly grade outcomes. A proposed action plan is presented to strengthen institutional support and refine the implementation of assessment strategies in Science classes.

[1] Lawigan-Bura National High School, Camiguin, Philippines
[2] Camiguin Polytechnic State College, Camiguin, Philippines
* joyna.jamero @deped.gov.ph

Volume 2, Issue 2, June 2026

Keywords: academic performance, authentic assessment, GUINSACAT

1. Introduction

In the global educational landscape, an ongoing paradigm shift has been endorsed toward 21st-century learning models that prioritize student-centered, active engagement over traditional rote memorization. This move aligns with the evolution of Sustainable Development Goal 4 (SDG 4) of the United Nations, which obligates every participating country to provide inclusive and equitable quality education. As required under Target 4.7, the education sector is obliged to equip learners with the scientific literacy and critical thinking skills to address existing complex global challenges (Tasquier, 2022). For educators, particularly science teachers, this mandate poses a unique challenge: moving beyond simple fact delivery to ensure students can apply complex concepts to real-world scenarios.

To achieve a collective global goal, authentic assessment (AA) has been deeply considered as the gold standard for evaluating higher-order skills. More than two decades ago, Gulikers et al. (2004) theorized that professional scientific practice and deep learning can occur through simulations, e-portfolios, and inquiry-based investigations. In the most recent study, Vlachopoulos (2025) emphasizes that innovative methodologies, including AI-enhanced tasks and game-based role-play, significantly boost student engagement and align science curricula with the demands of the modern workforce. Other studies also consistently showed that when educators transition from traditional testing to performance-based tasks, students demonstrate a marked improvement in academic achievement and a more positive attitude toward the learning area (Darling-Hammond et al., 2020).

In the Philippines, while the pivotal role of teachers is highly recognized in transitioning to performance-based assessments in reflecting genuine students' mastery of lessons through the Department of Education (DepEd) Order No. 8, s. 20215. However, despite these guidelines, many junior high school (JHS) Science classrooms remain heavily reliant on conventional paper-and-pencil tests. The result of this conservative approach is reflected in the assessments conducted by external organizations, such as the Trends in International Mathematics and Science Study (TIMSS) and the South East Asia Primary Learning Metrics (SEA-PLM). Results showed that Filipino students frequently struggle with items requiring application and critical reasoning (UNICEF, 2020). This clearly suggests that the reliance on traditional testing often fails to capture the true depth of a student’s understanding, potentially hindering their academic trajectory as scientific material becomes increasingly complex.

Despite these trends, there is a significant lack of empirical data within the local district on teachers' assessment practices for in-depth measurement of students' academic achievement. Although a plethora of broad studies relative to assessment exists, there is a dearth of evidence focused on the specific professional experiences

of JH Science teachers and the academic performance of their students in the local context. Grounded in Piaget's Constructivist Learning Theory (1954), which posits that students construct knowledge through meaningful interaction, this study aims to address the existing research gap. This study investigates the current extent of AA integration among JHS teachers and its effect on student academic performance during the first quarter of the 2025–2026 school year.

This research specifically aims to determine the level of AA integration among JHS science teachers and its relationship with students' academic performance. Furthermore, it seeks to identify if teacher demographics such as age, teaching experience, and educational attainment significantly influence these instructional choices. The findings are expected to serve as the foundation for an action plan to empower JHS teachers with effective, authentic assessment strategies that promote meaningful, lifelong scientific learning.

2. Methodology

2.1 Research Design

This study utilized a quantitative research approach and employed a descriptive-correlational design. This design was selected because it allows for the systematic description of existing phenomena and the examination of the nature and strength of relationships between variables, specifically the integration of authentic assessment and academic performance, without manipulating the research environment (Creswell & Creswell, 2023).

In a non-experimental design, no interventions or treatments were introduced during the study. Instead, data were collected as they naturally occurred in the classroom setting. This approach is particularly effective in educational research for identifying associations between pedagogical practices and student outcomes (Seeram, 2019). However, it is recognized that the descriptive-correlational nature of this study is constrained by the data sources utilized. Specifically, the use of teacher-reported practices paired with teacher-generated grades introduces a degree of circularity that may compromise the design's ability to detect a true objective relationship between the variables. To provide a deeper context for the quantitative findings, the survey results were supplemented with informal interviews, creating a more comprehensive profile of the assessment challenges and successes experienced by the JHS Science teachers.

2.2 Respondents and Sampling of the Study

The respondents comprised the total population of public Junior High School (JHS) Science teachers in the Guinsiliban, Sagay, and Catarman districts, from both integrated and secondary schools. To ensure respondents had sufficient contextual

knowledge of the local educational landscape and student demographics, an inclusion criterion of at least 1 year of teaching experience in SDO Camiguin was applied.

Universal sampling (complete enumeration) was employed to eliminate sampling bias and ensure the highest level of data representation. This technique involves studying the entire population of interest rather than a subset, which is particularly effective when the population size is manageable and specific characteristics are followed (Etikan et al., 2016). A total of 40 JHS science teachers comprised the respondents, distributed as follows: 15 teachers from the Catarman district, 15 from the Sagay district, and 10 from the Guinsiliban district. While this N=40 covers the entire target population in the identified districts, it is acknowledged that this sample size may be limited for establishing robust correlational relationships. Specifically, when respondents are partitioned into demographic sub-groups for comparative analysis, the resulting small and uneven cell sizes may reduce statistical power. This study explicitly acknowledges this sample size as a limitation in interpreting the results, particularly in the test of significance when the explanatory variable grouped respondents.

These schools are located in municipalities with active local economies and robust community stakeholder support, providing a stable environment for implementing performance-based assessment strategies. By including every eligible JHS science teacher in these districts, the study ensures that the findings comprehensively reflect the professional practices and challenges inherent in the local science curriculum during the 2025–2006 school year.

2.3 Research Instruments

This study utilized a structured research instrument adapted and modified from the works of Osin and Sahoo (2022) and Ozan (2019). Based on these previous studies, the survey questionnaire was designed to capture the nexus between performance-based tasks and the application of scientific concepts in real-world contexts. The modification of the tool was aligned with DepEd Order No. 8, s. 2015, which requires all public elementary and secondary schools to use the principles of performance-based evaluation to accurately reflect learners' progress within the K to 12 Science curriculum.

This instrument consisted of two sections. The first section (Part I) collected baseline data from the study's respondents. These data include the age, years of teaching experience, and highest educational attainment. Through this, more information can be generated about the integration of AA in the public JHS science classrooms.

The second part of the instrument contained the latent variable, the integration of AA in teaching science subjects. This includes 15 constructs or indicators specifically designed to evaluate the extent of AA integration in science instruction. These constructs covered various dimensions, including inquiry-based investigations, simulations, and e-portfolios. In quantifying the responses, a 4-point Likert Scale was

used, from “No Extent” (1) to “High Extent” (4), with descriptions “Not Integrated” to Extensively Integrated”. The use of a 4-point scale (forced-choice) is supported in educational research to eliminate neutral response bias and encourage definitive feedback from professional practitioners (Chomeya, 2010).

The study also used secondary data, namely students' grades or ratings from the first quarter of the school year 2025-2026. This served as the metric for academic performance, allowing for a correlational analysis between teacher pedagogical practices and student outcomes, as supported by the descriptive-correlational framework (Seeram, 2019). A critical limitation in this measurement approach is hereby acknowledged. The quarterly grades used as the dependent variable are generated by the teacher respondents themselves. Because these grades are a direct result of the individual teacher’s assessment methods and grading leniency, they may not provide an independent, objective measure of students' cognitive outcomes. The study recognizes that teacher-generated grades can be influenced by subjective factors, potentially resulting in a "ceiling effect" that can obscure the statistical relationship between integrating authentic assessment and actual student achievement.

2.4 Data Gathering Procedure

In gathering the required data, this inquiry had three (3) stages or processes to ensure rigor, transparency, and adherence to research protocols. In the first stage, it comprised the preliminaries of the inquiry. This included the school head's permission, followed by the requisition of authority to implement the study from the Office of the Schools Division Superintendent of DepEd-Camiguin. Upon approval, formal letters were sent to the Public Schools District Supervisors and School Heads of Guinsiliban, Sagay, and Catarman for their cooperation and support. These moves were carried out in accordance with the established research protocols of DepEd-Camiguin.

The second stage was about the actual implementation of the study. The researcher personally distributed the survey questionnaire to the 40 JHS science teachers. Creswell and Creswell (2023) confirmed that personal distribution is an effective strategy for ensuring high response rates and providing immediate clarification of complex indicators. Again, the researcher explained to the target respondents that their cooperation in completing the instrument constituted their consent. The researcher also requested in person that the instrument be returned after two weeks.

In the final stage, the researcher personally retrieved the survey questionnaire from the respondents. Before thanking the respondents, the researcher examined the instrument to determine whether any items were missing or blank. Additionally, a random, casual interview was conducted as a follow-up to the responses in the instrument. These data were tallied and encoded in the Excel format for analysis.

2.5 Data Analysis Procedure

After preparing the data in Excel format, it was exported to a CSV file for import into Jamovi, an open-source statistical software. The data analysis was carried out using the appropriate statistical tool to ensure reliable results. The assumptions of the parametric tests were thoroughly examined throughout the process. When one of the assumptions was not met, the alternative nonparametric test was used to analyze the data.

2.6 Ethical Considerations

To ensure the protection and well-being of all participants, this study strictly adhered to established ethical protocols of DepEd-Camiguin and the Data Privacy Act of 2012 (Republic Act No. 10173). Confidentiality and anonymity were maintained by using coded identifiers and by encrypting all digital data, ensuring that the research process posed no professional or personal risk to the educators involved. These precautions were fundamental in maintaining the study's integrity while respecting the fundamental rights and privacy of the academic community (Vlachopoulos, 2025).

3. Results

Table 1 presents the demographic profile of the study's respondents. It can be inferred that 18 (45%) of 40 respondents were in the early adulthood stage. While a quarter of them reached the late middle age stage, 12 (30%) belonged to the early middle age stage. The number of years of teaching showed that 22 (55%) of the respondents have served for more than 10 years. Only 6 (15%) worked for 1 to 6 years, while 12 (30%) reached 7 to 9 years of teaching experience. In terms of the educational attainment of the respondents, 13 (32.50%) of them finished at least a master's degree program in education, and only 8 (20%) earned units in a master's degree program in the graduate school. However, 19 (47.50%) have not yet engaged or enrolled in master's degree programs in any graduate school in the country.

Table 2 shows the extent of the integration of authentic assessment in teaching Science with 15 indicators grouped into four (4) constructs. It can be noticed that all indicators were rated with a qualitative description of "High Extent," leading to an overall mean (M=3.64) which is also at "High Extent". While the standard deviation (SD=0.24) indicates the similarities or consensus of the responses among the respondents of the study.

For the first construct, "Alignment and Cognitive Depth," the indicator with the highest mean (M = 3.98) indicates a commitment to aligning assessments with learning objectives and competencies. This suggests that the respondents prioritize instructional intent and cognitive rigor, emphasizing higher-order thinking skills (M = 3.70) and the practical application of knowledge (M = 3.65) over simple rote memorization. By ensuring assessments are purposeful and intellectually demanding,

Table 1. Demographic profile of the respondents of the study.

Variables	Counts	Percentage (%)
Age (years)		
22-34 (Early Adulthood Stage)	18	45.00
35-44 (Early Middle Age)	12	30.00
45-64 (Late Middle Age)	10	25.00
Number of Years in Teaching		
1-3	4	10.00
4-6	2	5.00
7-9	12	30.00
10 and above	22	55.00
Highest Educational Attainment		
Bachelor's Degree	19	47.50
With Units in Master's Degree	8	20.00
Completed Master's Degree	7	17.50
With Doctoral Units	6	15.00
Total	40	100.00

Table 2. Extent of the integration of authentic assessment in teaching science among JHS students.

Constructs/Indicators	Mean	Interpretation
I. Alignment & Cognitive Depth		
I align assessment with learning objectives and competencies.	3.98	High Extent
I assess higher-order thinking skills (e.g., analysis, evaluation).	3.70	High Extent
I assess application of knowledge, not just recall.	3.65	High Extent
II. Authentic & Performance-Based		
I emphasize both the learning process and the final output.	3.73	High Extent
I create tasks that mirror workplace or life situations.	3.70	High Extent
I use performance-based tasks (e.g., projects, experiments).	3.60	High Extent
I design/use real-world problems or scenarios in assessment tasks.	3.58	High Extent
I collect/gather multiple forms of student work as evidence (e.g., journals, portfolios).	3.58	High Extent
III. Feedback & Metacognition		
I involve students in self-assessment and reflection.	3.73	High Extent
I give timely and meaningful feedback.	3.73	High Extent
I use rubrics or clear criteria to assess outputs.	3.58	High Extent
I include peer assessment strategies.	3.43	High Extent
IV. Inclusivity & Learner Activity		
I adapt assessments to diverse learners' needs and contexts.	3.63	High Extent
I provide collaborative or group assessment tasks.	3.58	High Extent
I encourage student choice in task selection.	3.50	High Extent
Overall Mean	3.64	High Extent
Standard Deviation	0.24	

educators effectively bridge the gap between theoretical instruction and functional competence.

While the second construct on "Authentic and Performance-Based Assessment" clearly prefers valuing of the learning process alongside the final output (M = 3.73). This also shows a concerted effort to bridge the classroom and the

professional world, as evidenced by high scores in creating tasks that mirror workplace situations (3.70). While the use of portfolios and real-world scenarios remains high (3.58), these slightly lower means compared to objective alignment suggest that while the intent to be authentic is strong, the logistical implementation of complex, multi-modal evidence collection may be more challenging to execute consistently.

The third construct on “Feedback and Metacognition” reveals a high value placed on student reflection and timely intervention, both scoring (M = 3.73). This indicates a classroom environment that views assessment as a dialogue rather than a one-way delivery of grades. However, it is notable that "peer assessment strategies" yielded the lowest mean in the entire dataset (M = 3.43). This suggests that while instructors highly value their own feedback and the students’ self-reflection, they may be more cautious or less equipped to facilitate peer-to-peer evaluations, perhaps due to concerns regarding student objectivity or the complexity of moderating peer grades.

Finally, the data regarding “Inclusivity and Learner Activity” reflects a supportive learning environment, particularly in adapting assessments to diverse needs (3.63). While instructors show a strong inclination toward collaborative tasks (3.58), the practice of allowing student choice in task selection (3.50) is less frequent. This suggests a pedagogical model where the teacher remains the primary architect of the assessment structure. Overall, the findings portray a sophisticated assessment landscape deeply rooted in alignment and feedback, with minor opportunities for growth in delegating evaluative power to students through peer review and increased task autonomy.

Table 3 illustrates the academic performance of Junior High School students in Science, who were the respondents in the study. More than half of the student population reached the higher tiers of the grading scale, with 417 students (30.55%) achieving a "Very Satisfactory" rating and 385 students (28.21%) categorized as "Outstanding." Meanwhile, 333 students (24.39%) received a "Satisfactory" standing, and 198 students (14.51%) were classified as "Fairly Satisfactory." Notably, only a marginal group of 32 students (2.34%) fell into the "Did not meet expectations" category, indicating that while the majority are successful, a small percentage of the population still requires targeted academic support or remedial intervention.

Table 3. Academic performance of the JHS students in the science subject.

Grading Scale	Description	Counts	Percentage (%)
90-100	Outstanding	385	28.21
85-89	Very Satisfactory	417	30.55
80-84	Satisfactory	333	24.39
75-79	Fairly Satisfactory	198	14.51
Below 75	Did not meet the expectations	32	2.34
Total		1365	100.00
Average Grade = 85.33 (Very Satisfactory); SD= 1.87			

Overall, the proficiency of the students was pegged at the “Very Satisfactory” performance level, with a mean of 85.33 and a standard deviation of 1.87. It can be

said that there was a high level of consistency in student performance, reflecting a homogeneous distribution of grades, meaning the ratings clustered closely around the mean with very little variance. There was no wide gap between the highest and lowest-performing students in the group.

For the primary objective of this inquiry, the Spearman rank-order correlation was used as an appropriate statistical tool because the data were not normally distributed. The average of students' ratings per respondent (directly matched to survey respondents) was used as the academic performance paired with the perceived integration of authentic assessment. Table 4 shows that the analysis yielded a correlation coefficient of -0.006, indicating an extremely weak negative relationship between the two variables. The non-significant relationship does not conclusively demonstrate that AA has no impact on performance, but rather highlights the challenge of detecting relationships with aggregated quarterly grades that include multiple assessment types (written tests, quizzes, projects, participation) rather than AA-specific outputs.

Table 4. Test of significant relationship between the integration of authentic assessment and academic performance of JHS students in science.

Variables	r_s	r^2	p-value	df	Statistical Decision	Interpretation
Integration of Authentic Assessment vs Academic Performance	-0.006	0.000	0.971	38	Failed to reject Ho	Not Significant

Observing no normal distribution of data when grouped by variable, the Kruskal-Wallis H Test was considered the appropriate statistical tool for the test of significant difference. Table 5 shows that there were no significant differences in the extent of integration of authentic assessment among the respondents when grouped by age (H(2)=5.00,p= .82), number of years of teaching (H(3)=5.72,p= .126), and the highest educational attainment (H(3)=3.90,p= .273). This suggests that age, years of teaching experience, and teachers' highest educational attainment did not influence the degree of integration of authentic assessment in the Science class. However, these non-significant results must be interpreted with academic caution. As the reviewer noted, the distribution of respondents across demographic subgroups is highly uneven, with very small cell sizes (e.g., categories with only 2 or 4 respondents). Because the Kruskal-Wallis H Test lacks sufficient statistical power when applied to such small sub-samples, the findings are highly vulnerable to a Type II error—failing to detect a significant difference that might actually exist in the population. Consequently, while the statistical output suggests that adoption is not a function of maturity or tenure, the test's lack of sensitivity due to the small N may also contribute to these results.

Table 5. Test of significant difference on the extent of integration of authentic assessment grouped by variable.

Variables	H	df	p-value	Decision
Age	5.00	2	.082	Failed to reject Ho
Number of Years in Teaching	5.72	3	.126	Failed to reject Ho
Highest Educational Attainment	3.90	3	.273	Failed to reject Ho

4. Discussion

The demographic profile of the respondents shows a teaching force with maturity and professional stability; more than half are middle-aged, having served for more than 10 years. The respondents have typically reached high self-efficacy and pedagogical resilience (Viac & Fraser, 2020). However, a gap in formal academic advancement revealed that almost half of the respondents have not yet enrolled in graduate studies. This suggests that while the respondents possess significant practical experience, there is a need to bridge further the gap between long-term classroom practice and contemporary research-based instruction (Ibon et al., 2025). Locally, these findings mirror the career patterns in the Philippine education system, where a decade of service often marks a plateau in professional growth unless complemented by post-graduate schooling (Romorosa & Ucang, 2025). The high percentage of teachers yet to pursue advanced degrees aligns with studies that show administrative demands often hinder Filipino educators from pursuing higher academic qualifications (Chin, 2022). Consequently, while the workforce is stable and experienced, institutional support is necessary to incentivize graduate education and ensure that seasoned expertise is integrated with the latest scientific and pedagogical advancements.

The respondents of the study have demonstrated a "High Extent" of integration of authentic assessment. This implies a significant shift from the traditional approach toward a more holistic, competency-based approach. This suggests that educators are moving beyond rote memorization to ensure Science instruction bridges the gap between theory and functional competence. This is supported by Lund and Tannehill (2021), who emphasize that authentic assessment must meet the objective to ensure that the cognitive demand matches the intended learning outcomes of the 21st-century skills set. Sheik and Manap (2024) found in a study that Science teachers in the Philippines are increasingly using performance-based tasks to meet government-set standards.

Respondents also place a high value on student reflection and timely intervention, indicating a classroom environment that considers assessment as a dialogue. In the international context, Tutunaru (2023) argues that while teacher feedback is vital, it still requires student involvement as the center of the educative process. While Sabio and Sabio (2024) noted that Philippine educators often struggle with peer-led assessments due to cultural differences. This is due to the perceived authority of the teacher and concerns over students' success.

Finally, the construct of inclusivity and students' activities reflect a supportive environment, particularly for students' diverse needs. However, the respondents were cautious about granting full autonomy to the students' activities. In this regard, Sheikh and Manap (2024) suggested that assessment must be inclusive and authentic to promote students' capabilities. Locally, Ibon et al. (2025) highlighted that while Filipino teachers are highly empathetic toward diverse learners, logistical constraints and other factors affect students' success. In summary, these findings reveal an assessment approach deeply rooted in alignment and feedback, with significant opportunities to empower students through peer review and increased task autonomy.

The robust academic performance of the respondents' students in Science at "Very Satisfactory" performance could have been due to the alignment of the instructional delivery to the needs of the students. It is still too premature to say so. Further analysis is required for this. However, the academic performance can be sufficient to say that this is consistent with recent findings in the Philippines that the integration of technology in teaching Science can effectively enhance cognitive outcomes (Acut & Antonio, 2023). In other countries, positive results of a teaching approach are considered a byproduct (Jamaluddin et al., 2025). In terms of the distribution of grades, this indicates the narrowing of achievement gaps among the cohort. Magulod (2019) supported this finding by emphasizing that the consistency often stems from the prevalence of shared learning styles among Filipino students, backed by standardized materials. Similarly, Loyola-Carrillo et al. (2025) confirmed that a stable educational environment likely fosters performance clustering, as students benefit from collective momentum and a significant lack of pedagogical outliers. However, a subset of students still failed to meet expectations. While literature suggests that socio-economic stressors or a lack of grit might influence performance (Amistad-Sarzadill, 2025), these specific variables were not measured in the current study. Therefore, while these external factors remain plausible explanations for the observed academic gap, definitive conclusions about their impact on this specific cohort cannot be drawn without further empirical testing.

For the study's primary objective, the non-significant relationship between integrating authentic assessment and academic performance suggested that students' ratings were influenced by factors not accounted for by the researchers. However, it must be noted that the way the variables were measured may not be strong enough to support the claim of a relationship. The integration of authentic assessment was based on teachers' self-reports, while academic performance was based on students' first-quarter grades. These two data sources may not align directly because many factors, such as prior knowledge, grading practices, written work, performance tasks, attendance, teacher leniency, and school-level policies, can influence grades. Crucially, relying on teacher-generated quarterly grades as the metric for academic performance introduces circularity that likely obscures the true picture of academic achievement. Because these grades result directly from teachers' assessment methods, they may reflect teacher-specific grading rubrics rather than an independent, objective measure of cognitive gain. This methodological choice is very likely the reason why

the Spearman rank-order correlation yielded a coefficient of -0.006 and a p-value of 0.971, as teacher-generated grades often suffer from a "ceiling effect" that prevents the detection of a statistical relationship between pedagogical strategy and learning outcomes. Because of this, the finding of “no significant relationship” may not really mean that authentic assessment has no connection with performance; it may only show that the study design and measurement were not precise enough to detect the relationship.

If not, this could be due to the ceiling effect of both variables, where more than half of the students have at least a rating of 85 and similar responses at “High Extent” in the indicators of authentic assessment. In the local context, where the perceived integration of innovative practices does not translate into academic performance, Koh's (2019) study confirmed that this may be due to implementation challenges such as large class sizes or students' prior knowledge. Dubuc et al. (2022) and Lay (2022) also supported the observation that the non-translation of the intervention into students' ratings could be due to students’ internal factors, such as grit and prior knowledge, which are often considered more powerful predictors of Science achievement.

The test of difference showed no significant variations or influences in the extent of integration of authentic assessment among the respondents when grouped by their demographic variables. This implies that the adoption of strategies or innovations does not depend upon the professional maturity or advanced academic status of the teachers. Lay (2022) posited that teacher self-efficacy and individual technological pedagogical knowledge (TPACK) are far more influential than age or tenure in implementing innovative assessments. While Jamaluddin et al. (2025) confirmed that, in the Asia-Pacific region, when a curriculum is highly standardized, teachers tend to implement assessment tasks uniformly regardless of their length of service in the school. In the context of this work, the homogeneity in practice can be attributed to the standardized professional development programs, such as the Learning Action Cell (LAC) sessions, as mandated by the Department of Education, to ensure that all teachers, regardless of their educational background, are provided with the same training on assessment standards (Koh et al., 2019). The study by Acut and Antonio (2023) emphasized that institutional climate and the availability of shared resources level the playing field, leading to a uniform application of teaching innovations or strategies with the same objective: improving students' learning. However, it is also important to consider the statistical constraints of these findings. As noted, reliance on a total of 40 respondents resulted in very small and uneven cell sizes when grouped by demographic categories. This lack of statistical power makes the non-significant findings highly vulnerable to Type II error, in which existing differences in pedagogical practice might remain undetected due to the limited sensitivity of the Kruskal-Wallis H Test in such small samples. In this inquiry, this implies that while the integration of authentic assessment appears to be determined by systemic institutional requirements, future research with a larger sample size is necessary to confirm the absence of demographic influence.

5. Conclusion

The study concludes that the respondents are professionally stable and highly experienced Science teachers. However, a significant portion has not yet engaged in graduate-level studies, which remains a key area for professional advancement within the Department of Education. The JHS Science teachers demonstrated a high level of proficiency in integrating authentic assessments, specifically in aligning the cognitive domain with the required learning competencies of the K to 12 curriculum.

Regarding student outcomes, the learners demonstrated high performance at the "Very Satisfactory" level. However, the study revealed that the high level of integration of authentic assessment did not statistically affect students' academic achievement. In line with the reviewer's critique, the previous attribution of this result to "student grit and prior knowledge" has been removed, as these specific variables were not measured or tested within the scope of this research. While the descriptive data shows extensive implementation of authentic tasks, the Spearman rank-order correlation confirms that, for this specific cohort, there is no significant relationship between teacher-reported assessment practices and quarterly grades.

Finally, the high level of integration of authentic assessment did not vary among respondents, regardless of age, years of teaching, or highest educational attainment. This homogeneity can be attributed to the standardized professional development programs and institutional mandates provided by the government to ensure the uniform delivery of quality education. Nonetheless, it is acknowledged that the non-significant differences observed among demographic groups may also be due to limited statistical power resulting from small sub-group sizes, which increases the risk of a Type II error.

Acknowledgment

Sincere appreciation is given to all peer reviewers for their valuable comments and suggestions, which helped the authors to improve the quality of the manuscript.

Conflict of Interest Statement

The authors declare no conflicts of interest regarding the publication of this study.

References

Acut, D., & Antonio, R. (2023). Effectiveness of Science-Technology-Society (STS) approach on students' learning outcomes in science education: Evidence from a meta-analysis. *Journal of Technology and Science Education, 13*(3), 718. https://doi.org/10.3926/jotse.2151

Amistad-Sarzadilla, K. M. (2025). Grit Level and Academic Performance in Science of Students at Risk of Dropping Out. *Journal of Interdisciplinary Perspectives, 3*. https://doi.org/10.69569/jip.2024.0621

Chin, J. M.-C., Ching, G. S., del Castillo, F., Wen, T.-H., Huang, Y.-C., del Castillo, C. D., Gungon, J. L., & Trajera, S. M. (2022). Perspectives on the Barriers to and Needs of Teachers' Professional Development in the Philippines during COVID-19. *Sustainability, 14*(1), 470. https://doi.org/10.3390/su14010470

Chomeya, R. (2010). Quality of Psychology Test Between Likert Scale 5 and 6 Points. *Journal of Social Sciences, 6*(3), 399-403. https://doi.org/10.3844/jssp.2010.399.403

Creswell, J. W., & Creswell, J. D. (2023). Research design: Qualitative, quantitative, and mixed methods approaches (6th ed.). SAGE Publications. https://books.google.com/books?hl=en&lr=&id=335ZDwAAQBAJ&oi=fnd&pg=PT16&dq=Creswell,+J.+W.,+%26+Creswell,+J.+D.+(2023).+Research+design:+Qualitative,+quantitative,+and+mixed+methods+approaches+(6th+ed.).+SAGE+Publications.+&ots=YEAPHItovL&sig=PuFhw3a7KSiF-BxCKV02lqEJyP0

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24*(2), 97-140. https://doi.org/10.1080/10888691.2018.1537791

Department of Education. (2015). Policy guidelines on classroom assessment for the K to 12 Basic Education Program (DepEd Order No. 8, s. 2015). https://www.deped.gov.ph/wp-content/uploads/2015/04/DO_s2015_08.pdf

Dubuc, M. M., Aubertin-Leheudre, M., & Karelis, A. D. (2022). Predictors of Academic Performance in High School Students: The Longitudinal ASAP Study. *International journal of exercise science, 15*(4), 616–631. https://doi.org/10.70252/ZAZQ9758

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11

Gulikers, J. T., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*(3), 67-86. https://doi.org/10.1007/BF02504676

Ibon, A. S., Pagayonan, E. N., & Galay-Limos, J. A. (2025). Innovative teaching strategies and adaptations of science teachers in secondary schools of Sablayan District. *International Journal of Research Studies in Management, 13*(8), 75-92. https://doi.org/10.5861/ijrsm.2025.25522

Jamaluddin, F., Razak, A. Z. A., & Rahim, S. S. A. (2025). Navigating the challenges and future pathways of STEM education in Asia-Pacific region: A comprehensive scoping review. *STEM Education, 5*, 53-88. https://doi.org/10.3934/steme.2025004

Koh, K., Delanoy, N., Thomas, C., Bene, R., Chapman, O., Turner, J., ... & Hone, G. (2019). The role of authentic assessment tasks in problem-based learning. *Papers on postsecondary learning and*

teaching, 3, 17-24. https://journalhosting.ucalgary.ca/index.php/pplt/article/download/53144/51795/

Lay, A. A. M. (2021). The relationship between teachers' efficacy and classroom management. *Journal of Education and Practice, 12*, 26-33.

Loyola-Carrillo, P., Vega-Muñoz, A., Salazar-Sepúlveda, G., Gil-Marín, M., & Adsuar-Sala, J. (2025). Studying engagement in educational settings: a mapping review on high-impact academic engagement research. *Frontiers in Psychology, 16*. https://doi.org/10.3389/fpsyg.2025.1519509

Lund, J., & Tannehill, D. (2021). Standards-based physical education curriculum development. Jones & Bartlett Learning. https://books.google.com/books?hl=en&lr=&id=HYmOAwAAQBAJ&oi=fnd&pg=PR1&dq=Lund,+J.,+%26+Tannehill,+D.+(2021).+Standards-based+physical+education+curriculum+development.+Jones+%26+Bartlett+Learning.&ots=eXgaIZfzpS&sig=S9oGBvovMz_-pzF9UBhToEtnpnc

Magulod, G. (2019). Learning styles, study habits, and academic performance of Filipino University students in applied science courses: Implications for instruction. *Journal of Technology and Science Education, 9*(2), 184. https://doi.org/10.3926/jotse.504

Osin & Sahoo, M. (2022). Performance-based assessment in science at elementary school level. *International Journal of Advance and Applied Research, 9*(5), 646-650. https://mitanjalisahoo.github.io/publications/publication5.html

Ozan, C. (2019). Authentic Assessment Increased Academic Achievement and Attitude towards the Educational Measurement of Prospective Teachers. *International Journal of Evaluation and Research in Education, 8*(2), 299-312. https://eric.ed.gov/?id=EJ1220891

Piaget, J. (1954). The construction of reality in the child. New York, NY: Basic Book. https://scholar.archive.org/work/tgee3gmbxvga5pskmhu4r6xblq/access/wayback/http://pages.uoregon.edu/rosem/Timeline_files/The%20Construction%20of%20Reality%20in%20the%20Child.pdf

Romorosa, Q., & Ucang, J. (2025). Professional Development in the Lens of Public School Teachers. *International Journal of Research in Education and Science, 11*(1), 140-149. https://files.eric.ed.gov/fulltext/EJ1465829.pdf

Sabio, R. A., & Sabio, C. J. (2024). Challenges in Implementing Online Learning in Philippine Higher Education: Business Students' Perspective. *International Journal of Learning, Teaching and Educational Research, 20*(7), 262-290. https://www.ijlt.org/articles/2024/IJLT-V10N3-348.pdf

Seeram, E. (2019). An overview of correlational research. Radiologic Technology, 91(2), 176-179. https://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=00338397&asa=N&AN=139298367&h=0%2BXgdMlU%2BxweRuq7FdQZ1FV9hUMgmzSIZ%2FZeDngpqD98%2Fle44jZQ1tUInp%2BF6ajI0Vbh6LZsV3RmKyZPrws64w%3D%3D&crl=c

Sheikh, A. B. M., & Manap, M. N. A. (2024). Classroom Assessment Practices. *Open Journal of Social Sciences, 12*, 239-258. https://doi.org/10.4236/jss.2024.123018

Tasquier, G., Knain, E., & Jornet, A. (2022). Scientific literacies for change making: Equipping the young to tackle current societal challenges. *Frontiers in Education, 7*. https://doi.org/10.3389/feduc.2022.689329

Tutunaru, T. (2023). Improving assessment and feedback in the learning process: directions and best practices. *Research and Education, 8*, 38-60. https://researchandeducation.ro/2023/07/31/improving-assessment-and-feedback-in-the-learning-process-directions-and-best-practices.html

UNICEF. (2020). SEA-PLM 2019 main regional report: Children's learning in 6 southeast asian countries. https://research.acer.edu.au/ar_misc/52/

Viac, C., & Fraser, P. (2020). Teachers' well-being: A framework for data collection and analysis. *OECD Education Working Papers, 213*, 1-81. https://search.proquest.com/openview/4a2fab8930833e04411ce2583993c8ab/1?pq-origsite=gscholar&cbl=54479

Vlachopoulos, D. (2025). Authentic Assessment: Bridging Higher Education And Real-World Skills Through Diverse Methodologies. In INTED2025 Proceedings (pp. 2225-2234). IATED. https://library.iated.org/view/VLACHOPOULOS2025AUN

Author Contributions: Jamero, J., Sabellina, J.; Study design, method conception, data collection, data analysis and manuscript writing.