

**Performance-based assessment in Araling Panlipunan 7: Examining utilization,
instructional effectiveness, and implementation challenges
in a Philippine District context**

Deanne W. Baliwag*

Kalinga State University, Tabuk City, Kalinga, Philippines

ABSTRACT

This study investigated the utilization, effectiveness, and implementation challenges of performance tasks in teaching Araling Panlipunan 7 within selected districts in the Philippines. Anchored in constructivist learning theory, this research responds to the increasing demand for learner-centered pedagogies, contextual understanding, and real-world application. Findings revealed that the overall extent of use of performance tasks yielded an average weighted mean of 3.32, interpreted as “sometimes,” indicating moderate yet inconsistent integration of performance-based assessment strategies. In terms of effectiveness, performance tasks were perceived as “much effective” (overall mean = 3.54), indicating that students recognize their value in enhancing engagement, critical thinking, and conceptual understanding. High-performing tasks demonstrated strong alignment with authentic assessment principles, particularly in fostering creativity and contextual learning. The analysis of challenges encountered by teachers yielded an overall mean of 2.47, interpreted as “slightly serious.” While material and infrastructural constraints were not perceived as major barriers, issues related to teacher capacity - particularly limited research skills and insufficient training opportunities - emerged as moderately serious concerns. These findings align with SDG 4 on Quality Education where performance tasks in Araling Panlipunan 7 are generally perceived as effective and are moderately utilized, addressing challenges requires a systemic approach that integrates policy, practice, and professional development to fully realize the transformative potential of authentic assessment in enhancing student learning outcomes. Its sustainability contribution lies in fostering collaboration among Araling Panlipunan teachers to improve learner engagement and outcomes.

Keywords: Performance-based assessment, Araling Panlipunan, constructivist learning theory.

Date Submitted: May 1, 2026

Date Accepted: May 6, 2026

Date Published: May 13, 2026

INTRODUCTION

The increasing emphasis on learner-centered pedagogy and authentic assessment has fundamentally transformed contemporary educational practice, particularly in the context of 21st-century skills development. Across global education systems, performance-based

*Corresponding author / Email: deannebaliwag0@gmail.com

DOI: <http://doi.org/10.69651/PIJHSS05021166>

Recommended citation:

Baliwag, D. W. (2026). Performance-based assessment in Araling Panlipunan 7: Examining utilization, instructional effectiveness, and implementation challenges in a Philippine district context. *Pantao (International Journal of the Humanities and Social Sciences)* 5 (2), 4503-4521. <http://doi.org/10.69651/PIJHSS05021166>

assessment has emerged as a critical mechanism for fostering higher-order thinking, problem-solving, and contextualized learning (Koh & Fu, 2025; Rehman et al., 2023). Unlike traditional assessment approaches that prioritize recall and discrete knowledge acquisition, performance tasks require students to actively construct meaning through real-world applications, thereby aligning with constructivist and sociocultural learning theories (Deane, 2023; Wood, 2025). In the field of Social Studies education, particularly in Araling Panlipunan, such approaches are essential for cultivating historical thinking, civic competence, and critical awareness of societal issues (Fang, 2023; Beach et al., 2022). This theoretical grounding positions performance-based assessment not merely as an evaluative tool but as a pedagogical instrument that supports the holistic development of learners within meaningful and contextually relevant educational experiences.

In the Philippine educational context, the Department of Education has institutionalized performance tasks as a core component of classroom assessment under the Enhanced K to 10 curriculum. Rooted in the foundational vision articulated by Gregorio (1976), that education in the Philippines is aimed at the development of the whole individual mentally, physically, socially, emotionally, and spiritually through interaction of experiences at home and in school, the curriculum envisions learning as an integrative and experiential process that extends beyond the confines of traditional classroom settings. This philosophy is operationalized through DepEd Order No. 8, s. 2015, which provides the policy guidelines on classroom assessment for the K to 12 Basic Education Program and explicitly positions performance tasks as a key component of both formative and summative assessment across all subject areas, including Araling Panlipunan for Grade 7. Complementing this, DepEd Order No. 21, s. 2019 specifies that performance tasks must reflect real-life application of knowledge, while DepEd Order No. 42, s. 2016 mandates the use of lesson exemplars in daily instruction to guide teachers in the design and delivery of these tasks. The broader legislative framework is anchored in Republic Act No. 10533, otherwise known as the Enhanced Basic Education Act of 2013, which establishes the K to 12 curriculum as the foundation for ensuring every child's access to quality, globally competitive education. More recently, the DepEd MATATAG K-10 Curriculum has further reinforced these commitments by centering learner-centered, competency-based, and authentic assessments as the cornerstones of instructional practice, with performance tasks in Araling Panlipunan 7 designed to assess students' understanding and application of concepts through diverse activities that go beyond written examinations.

Despite this robust policy architecture, empirical evidence consistently points to a persistent gap between the intended and actual implementation of performance-based assessment in classrooms (Alghammas, 2020; Sudimantara et al., 2025). This gap is frequently attributed to contextual constraints including limited teacher training, inadequate instructional resources, and competing curricular demands (Ma'rufah et al., 2024; Zaabalawi & Zaabalawi, 2024). The effectiveness of performance tasks, as the literature affirms, is contingent upon multiple interrelated factors including teacher assessment literacy, instructional design, and institutional support systems (Koh & Fu, 2025; Lin et al., 2025). Studies have shown that while teachers generally recognize the pedagogical value of authentic assessment, they often encounter difficulties in designing tasks that are both cognitively demanding and aligned with learning objectives (Puryanto et al., 2021; Murray & Christison, 2020). The integration of multimodal and technology-enhanced performance tasks introduces additional complexities, particularly in under-resourced educational settings (Sudimantara et al., 2025; Lin et al., 2025). These challenges are further compounded in decentralized education systems such as the Philippines, where contextual factors including district affiliation, school leadership, and community support significantly influence the extent and quality of performance task utilization (Sudimantara et al., 2025; Wood, 2025).

Teacher preparedness represents a particularly consequential variable in this landscape. Padillo et al. (2021) highlight the importance of providing teachers with adequate training, resources, and professional development opportunities to effectively navigate curriculum changes, and this observation is especially salient under the MATATAG Curriculum, where successful implementation is essential to ensuring that all students are holistically developed and that the curriculum's goals of relevance, effectiveness, and responsiveness to diverse student needs are fully realized. Challenges in implementation often extend beyond resource limitations to include time management, workload, and insufficient professional development (Ma'rufah et al., 2024; Alghammas, 2020), all of which compound the difficulties teachers face in consistently and effectively deploying performance-based assessment strategies.

In Social Studies education specifically, performance tasks such as debates, role-playing, and document analysis hold particular value in promoting disciplinary literacy and critical inquiry (Wood, 2025; Asakavičiūtė et al., 2023). These tasks enable learners to engage with historical and contemporary issues in meaningful and sustained ways, fostering deeper understanding and long-term retention. However, the consistent implementation of such tasks in district-based and rural Philippine school settings remains insufficiently documented in the literature. Existing research has predominantly concentrated on language and STEM education, leaving a notable gap with respect to performance-based assessment in Araling Panlipunan (Rehman et al., 2023; Ramadhan et al., 2023). This disciplinary and contextual gap limits the available evidence base that policymakers and practitioners can draw upon when seeking to improve assessment practice within the Social Studies domain.

Student perceptions constitute another critical dimension through which the effectiveness of performance tasks must be evaluated. As primary stakeholders in the learning process, students offer invaluable insights into how performance tasks influence engagement, comprehension, and skill development. Research indicates that positive student perceptions are associated with increased motivation and improved learning outcomes (Biju et al., 2024; Lawrence & Sherry, 2021). Conversely, inconsistencies in task implementation may generate variability in perceived effectiveness, reflecting deeper disparities in instructional quality and classroom experience (Zaabalawi & Zaabalawi, 2024; Deane, 2023). Integrating the student perspective into the analysis of performance-based assessment thus provides a more complete and balanced account of its educational impact.

The conceptual framework guiding this study reflects these converging theoretical and empirical considerations. As illustrated in Figure 1, the independent variables encompass the frequency of performance task use, their perceived effectiveness, and the specific implementation challenges encountered by teachers in Grade 7 Araling Panlipunan. These variables are analyzed in relation to the dependent variable, which is the overall profile of performance task utilization, effectiveness, and challenges within the subject. The framework further incorporates moderating variables including sex, district affiliation, and access to technology, recognizing that these factors may shape how teachers and students experience and evaluate performance tasks. By accounting for these moderating influences, the study enables a more nuanced and contextualized understanding of the dynamics that determine the implementation and outcomes of performance-based assessment.

This study addresses the identified gaps in the literature by examining the utilization, effectiveness, and challenges of performance tasks in teaching Araling Panlipunan 7 within selected districts in the Philippines. Adopting a descriptive-analytical approach, it contributes empirical evidence to the ongoing discourse on authentic assessment and offers contextually grounded insights that can inform both policy and professional practice. The significance of this inquiry lies in its potential to bridge the distance between theoretical frameworks and

classroom realities, providing a foundation for the development of more effective assessment strategies and targeted professional development programs for Araling Panlipunan teachers. Ultimately, this study affirms the centrality of performance-based assessment in advancing the goals of 21st-century education, where learners are expected to demonstrate not only knowledge but also the capacity to apply, analyze, and create in meaningful and authentic contexts (Koh & Fu, 2025; Rehman et al., 2023), while simultaneously acknowledging the systemic conditions that must be addressed to fully realize this transformative potential.

Statement of the problem

This study looked into the extent of use, effectiveness, and challenges encountered regarding performance tasks in teaching Araling Panlipunan 7.

1. What is the extent of use of the performance tasks used by teachers in teaching Araling Panlipunan 7 as perceived by the students, and are there significant differences in the perceptions of the respondents as to the moderator variables of sex, district affiliation, and access to technology?
2. What is the level of effectiveness of the performance tasks used by teachers in teaching Araling Panlipunan 7 as perceived by the students, and are there significant differences in the perceptions of the respondents as to the moderator variables of sex, district affiliation, and access to technology?
3. What are the challenges encountered by teachers in the use of performance tasks in teaching Araling Panlipunan 7, and are there significant differences in the perceptions of the respondents on these challenges as to the moderator variables of sex, district affiliation, and access to technology?

METHODOLOGY

This chapter presents the methodological framework of the study, encompassing the research design, locale and population, data gathering instruments and procedures, ethical considerations, and the statistical tools employed in data analysis.

The study employed a descriptive quantitative research design, utilizing a survey questionnaire as the primary instrument for data gathering. According to Siedlecki (2020), descriptive research aims to measure and examine a unit, population, or situation and collect data on individuals or groups of individuals. This approach was appropriate as it enabled the researcher to determine the extent of use and assess the level of effectiveness of performance tasks in Araling Panlipunan 7, as well as to identify the challenges encountered by teachers during implementation.

The study was conducted in public junior high schools across the districts of Lubuagan, Pasil, and Tanudan within the Kalinga Division. Private schools and other districts were intentionally excluded to maintain a clear focus. The respondents were limited to Grade 7 students and Araling Panlipunan 7 teachers; learners from other grade levels, subject areas, or school roles were not included. The study focused on experiences and viewpoints during School Year 2025 to 2026 and was based on self-reported data, which may reflect individual perceptions and recollections. Academic achievement through standardized tests was not measured; instead, the study concentrated on reported experiences with performance-based assessment. These boundaries were established to allow an in-depth analysis of the selected context and to provide relevant recommendations for educational practices across the three districts in the Kalinga Division.

The total population of the study comprised 301 respondents distributed across nine schools in the Kalinga Division, consisting of 14 teachers and 287 Grade 7 students. Lubuagan National High School had the highest representation at 18.60% of the total respondents, with a frequency of 56, comprising 2 teachers and 54 students. Pasil National High School contributed 39 respondents, accounting for 12.96%, with 2 teachers and 37 students. Magtoma Pangol National High School had 38 respondents representing 12.62%, consisting of 1 teacher and 37 students. Lubuagan National High School Poblacion Extension contributed 33 respondents at 10.96%, with 1 teacher and 32 students. Biga National High School had 34 respondents representing 11.30%, composed of 2 teachers and 32 students. Tanudan National High School contributed 32 respondents at 10.63%, with 2 teachers and 30 students, while Batong Buhay National High School and Collayo Extension similarly contributed 32 respondents at 10.63%, also with 2 teachers and 30 students. Tanudan Vocational School had 22 respondents representing 7.31%, consisting of 1 teacher and 21 students. Taloctoc General Comprehensive National High School had the fewest respondents at 4.98%, with a frequency of 15 composed of 1 teacher and 14 students. The seven remaining schools showed a relatively even distribution, each contributing between approximately 7% and 13% of the total respondents.

Among the 287 student respondents, the sex distribution was relatively balanced, with 156 males comprising 54.4% and 131 females comprising 45.6% of the total. This balance strengthens internal validity, particularly in studies examining assessment practices where gendered learning preferences may influence responses, and minimizes gender bias in perception-based findings. The absence of significant variation across sex in subsequent analyses indicates that performance task utilization transcends gendered differences, aligning with contemporary assessment literature that emphasizes inclusivity in authentic assessment practices grounded in multimodal and participatory approaches.

In terms of district affiliation, student respondents were drawn from Tanudan (134 respondents, 46.7%), Lubuagan (86 respondents, 30.0%), and Pasil (67 respondents, 23.3%), reflecting a geographically diverse sample within a rural Philippine context. The prominence of Tanudan respondents suggests a potential weighting toward its contextual realities, including access disparities and localized teaching strategies. This distribution is significant, as contextual variability in educational resources and pedagogical practices often shapes instructional implementation. Subsequent significant differences across districts underscore the role of contextual factors in shaping performance task utilization, consistent with literature emphasizing that instructional innovation is mediated by localized conditions including infrastructure and institutional support, particularly in under-resourced settings.

With respect to access to technology, all 287 student respondents (100%) reported having access to technology, indicating a digitally enabled learning environment across the study locale. Because access was consistent across all respondents, it does not introduce unintended gaps or alter results when comparing the frequency and effectiveness of performance tasks. This uniformity also affirms that basic digital tools and facilities are now widely available in these districts and no longer constitute a limiting factor in the implementation of performance-based assessments. This finding supports emerging evidence that technology integration enhances authentic assessment practices by enabling multimodal tasks and interactive learning experiences, and suggests that pedagogical variation is more likely attributable to teacher competence and contextual factors rather than technological constraints.

Among the 14 teacher respondents, the sex distribution was predominantly female, with 11 female teachers comprising 78.6% and 3 male teachers comprising 21.4% of the group. This reflects the gendered composition of the teaching workforce in basic education. Existing

literature suggests that gender does not significantly influence pedagogical effectiveness in implementing performance-based assessments, and that competence, training, and instructional support are more critical determinants. By district affiliation, teacher respondents were distributed across Tanudan (7 teachers, 50.0%), Pasil (4 teachers, 28.6%), and Lubuagan (3 teachers, 21.4%). The concentration of teacher respondents in Tanudan suggests that findings may be more reflective of its contextual conditions, including institutional support and prevailing instructional practices. This distribution reinforces the argument that contextual and institutional factors shape pedagogical implementation, particularly in decentralized systems, consistent with literature affirming that localized conditions including leadership support and resource allocation critically influence the success of authentic assessment practices (Sudimantara et al., 2025; Alghammas, 2020).

The research instrument consisted of a structured survey questionnaire organized into three sections corresponding to the three research questions. The first section, composed of 20 items, was adapted from the Lesson Exemplars of the MATATAG K-10 Curriculum for Araling Panlipunan 7 covering Quarters 1 to 4, and was designed to determine the extent of use of performance tasks as perceived by students. The second section, similarly comprising 20 items sourced from the same MATATAG curriculum Lesson Exemplars, assessed the perceived level of effectiveness of the performance tasks implemented in the classroom. The third section, consisting of 7 items, was adapted from the study conducted by Margie V. Galvez (2024) and was designed to identify and understand the challenges teachers face when using performance tasks in Araling Panlipunan 7.

To establish content validity, the instrument was carefully designed in alignment with the DepEd MATATAG K-10 Curriculum, relevant official department guidelines, and established educational research, ensuring that every item corresponded precisely with the content, key skills, and required learning standards for Araling Panlipunan 7. The draft instrument underwent a rigorous content validation process, having been thoroughly reviewed by the research adviser and thesis panel members during the proposal defense. Reviewers assessed whether all items were clear, relevant, and correctly measured the intended constructs covering the extent of use, effectiveness, and challenges of performance tasks. All feedback and recommendations were fully incorporated to refine the instrument, ensuring accuracy, clarity, and readiness for actual data gathering.

Reliability was assessed using Cronbach's Alpha, the standard statistical method for surveys employing rating scales. This calculation was performed separately for each section of the questionnaire using actual responses from the study's participants. Data from 287 students were used to test reliability for the sections on extent of use and effectiveness of performance tasks, while responses from 14 teachers were used for the section on challenges encountered. The section on Extent of Use obtained a Cronbach's Alpha score of 0.89, the section on Effectiveness obtained 0.90, and the section on Challenges Encountered obtained 0.81. All values exceeded the widely accepted minimum benchmark of 0.70 established by Nunnally (1978), confirming that every section of the instrument demonstrated excellent to very high internal consistency. It is noted that while 14 teachers may appear to be a small sample, this number represents the total population of Araling Panlipunan 7 teachers in the covered areas, and the resulting reliability scores nonetheless meet and exceed the required standard, further affirming the instrument's dependability.

Before the study was conducted, the researcher secured approval of the research proposal from the proper institutional authorities. Formal permission was subsequently requested and obtained from the Schools Division Superintendent of the Kalinga Division to conduct the study across the districts of Tanudan, Pasil, and Lubuagan. Following the grant of permission, survey questionnaires were administered to the respondents, who were provided

with informed consent forms explaining that their participation was voluntary and confidential. Respondents were given sufficient time to complete the questionnaire, after which their responses were collected, tabulated, and analyzed in accordance with the objectives of the study.

The data gathered were analyzed using both descriptive and inferential statistical tools. Frequency count and percentage were employed to describe the socio-demographic profile of respondents with respect to sex, district affiliation, and access to technology, providing a clear description of the distribution across the sample. The weighted mean and standard deviation were used to determine the extent of use, level of effectiveness, and challenges encountered in the implementation of performance tasks in Araling Panlipunan 7. The Kruskal-Wallis H test, a nonparametric statistical procedure, was used to determine whether significant differences existed in respondents' perceptions when grouped according to sex, district affiliation, and access to technology. This test was selected because the data were derived from ordinal Likert scale responses and did not assume normal distribution. The level of significance was set at $p = 0.05$. The decision rule was as follows: if the p-value is less than or equal to 0.05, the null hypothesis is rejected and the result is interpreted as significant; if the p-value is greater than 0.05, the null hypothesis fails to be rejected and the result is interpreted as not significant.

The interpretation of data was guided by three distinct scoring scales corresponding to each research question. To determine the extent of use of performance tasks as perceived by students, a five-point Likert scale was applied with the following intervals and descriptive equivalents: a mean range of 4.21 to 5.00 corresponds to a Likert scale value of 5 and is described as "Always," indicating that performance tasks are consistently and systematically integrated in Araling Panlipunan 7 instruction, reflecting a high level of pedagogical alignment with authentic assessment frameworks and routine employment of performance-based strategies to facilitate higher-order thinking and contextual learning; a range of 3.41 to 4.20 corresponds to a value of 4 and is described as "Often," indicating that performance tasks are frequently utilized in classroom instruction, with teachers demonstrating strong adoption of performance-based approaches, although implementation may vary depending on lesson objectives or contextual constraints; a range of 2.61 to 3.40 corresponds to a value of 3 and is described as "Sometimes," indicating moderate and inconsistent integration reflecting transitional adoption of authentic assessment strategies contingent on teacher readiness, time constraints, or resource availability; a range of 1.81 to 2.60 corresponds to a value of 2 and is described as "Rarely," indicating infrequent use suggestive of barriers such as insufficient training, lack of instructional materials, or competing curricular demands; and a range of 1.00 to 1.80 corresponds to a value of 1 and is described as "Never," reflecting reliance on traditional assessment methods and minimal alignment with constructivist and performance-based assessment paradigms.

To determine the level of effectiveness of performance tasks as perceived by students, the same five-point Likert scale structure was applied with parallel descriptive equivalents contextualized to effectiveness: a mean range of 4.21 to 5.00 with a value of 5 is described as "Always," indicating that performance tasks are highly effective in consistently promoting higher-order thinking, critical analysis, and contextual understanding with strong alignment to authentic assessment and constructivist principles; a range of 3.41 to 4.20 with a value of 4 is described as "Often," indicating generally effective tasks that frequently support engagement and comprehension, though effectiveness may vary with implementation quality, teacher expertise, and resource availability; a range of 2.61 to 3.40 with a value of 3 is described as "Sometimes," indicating moderate and inconsistent effectiveness suggesting partial alignment with performance-based practices and possible limitations in execution or instructional design;

a range of 1.81 to 2.60 with a value of 2 is described as "Rarely," indicating limited effectiveness suggesting challenges in instructional delivery or assessment design and a lack of depth or alignment with learning objectives; and a range of 1.00 to 1.80 with a value of 1 is described as "Never," reflecting a disconnect from authentic assessment principles and reliance on ineffective or poorly designed assessment practices.

To identify the challenges encountered by teachers, a separate five-point Likert scale was applied with the following descriptive equivalents: a mean range of 1.00 to 1.80 with a value of 1 is described as "Not Serious"; a range of 1.81 to 2.60 with a value of 2 is described as "Slightly Serious"; a range of 2.61 to 3.40 with a value of 3 is described as "Moderately Serious"; a range of 3.41 to 4.20 with a value of 4 is described as "Serious"; and a range of 4.21 to 5.00 with a value of 5 is described as "Very Serious." The combination of these statistical tools allowed for a comprehensive analysis of how the independent variables influence the dependent variable in accordance with the study's conceptual framework.

RESULTS AND DISCUSSION

This chapter presents the results and discussion of the study, which examined the utilization, effectiveness, and implementation challenges of performance tasks in teaching Araling Panlipunan 7. The study involved 301 respondents composed of 287 Grade 7 students and 14 Araling Panlipunan 7 teachers drawn from nine public junior high schools across the districts of Tanudan, Pasil, and Lubuagan in the Kalinga Division. A descriptive quantitative research design was employed, utilizing a structured survey questionnaire adapted from the MATATAG K-10 Curriculum Lesson Exemplars and from the study of Galvez (2024). Data were analyzed using the weighted mean, standard deviation, and the Kruskal-Wallis H test at a 0.05 level of significance. The results are interpreted in direct relation to the three research objectives of the study and are discussed in conjunction with relevant scholarly literature to contextualize the findings within broader theoretical and empirical frameworks.

Extent of use of performance tasks in Araling Panlipunan 7

The first objective of the study sought to determine the extent to which performance tasks are utilized in teaching Araling Panlipunan 7 as perceived by the 287 student respondents. The overall average weighted mean of 3.32, falling within the interval of 2.61 to 3.40, yields a descriptive interpretation of "Sometimes," indicating that performance tasks are moderately but inconsistently integrated into classroom instruction across the three districts. This result reflects a transitional stage between traditional and authentic assessment practices, wherein teachers have begun incorporating performance-based strategies but have not yet achieved systematic or sustained adoption.

An examination of individual indicators reveals notable variability in the frequency of implementation across the 20 performance tasks assessed. Slogan Making obtained the highest mean of 3.86 (SD = 1.23), followed by Guess the Country or Place (M = 3.76, SD = 1.06), Sanaysay (M = 3.75, SD = 1.17), and Suri-Larawan (M = 3.72, SD = 1.12), all of which fall within the "Often" range of 3.41 to 4.20. Other tasks rated as "Often" include MaPanuri (M = 3.67, SD = 1.33), Hula Salita (M = 3.67, SD = 1.22), Kinig-Kanta (M = 3.59, SD = 1.08), Hula Bandila (M = 3.55, SD = 1.26), MAPA-Hanap (M = 3.47, SD = 1.37), Loop a Word (M = 3.52, SD = 1.26), and Suri-Simbolo (M = 3.43, SD = 1.31). When asked why these activities were most preferred and frequently used, students shared that they found Slogan Making enjoyable, straightforward, and expressive, allowing them to communicate what they learned creatively without feeling excessive pressure. They also noted that activities like Guess the Country,

Sanaysay, and Suri-Larawan were simple, interesting, and effective in making lessons easier to understand and remember. Teachers corroborated this by explaining that these activities fit well within the one-hour class schedule, are well-structured, and clearly reveal whether students have understood the lesson objectives. These tasks inherently promote higher-order thinking, creativity, and contextual analysis, which are core features of authentic assessment frameworks. Prior studies affirm that performance-based tasks, particularly those involving creative expression and inquiry, enhance student engagement and disciplinary understanding by situating learning in meaningful contexts (Koh & Fu, 2025; Rehman et al., 2023).

In contrast, several tasks fell within the lower range of the "Sometimes" category, approaching "Rarely" utilization. Freedom Wall obtained a mean of 2.76 (SD = 1.36), Panayam o Talkshow a mean of 2.74 (SD = 1.31), Suri-Balita a mean of 2.72 (SD = 1.23), and Imperialism Profile a mean of 2.80 (SD = 1.49). Additional tasks rated as "Sometimes" include Emoticon Decipher (M = 2.94, SD = 1.32), Founders Keeper (M = 3.05, SD = 1.37), Gallery Walk (M = 3.02, SD = 1.14), Debate (M = 3.10, SD = 1.32), and Ikuwento, Wariin, Hanapin, Lagumin or K-W-H-L Chart (M = 3.36, SD = 1.39). Students explained that activities such as Panayam o Talkshow, Suri-Balita, and Imperialism Profile require extensive time, thorough preparation, and considerable effort, and that some students feel nervous or anxious when required to speak or perform in front of the class. Teachers acknowledged the value of these tasks while noting that they find them challenging to facilitate and manage. Collaborative and interactive strategies such as Debate and Gallery Walk were similarly constrained, with students noting that these activities require substantial space, longer time allocations, and careful planning that are difficult to accommodate within a regular one-hour class schedule. Teachers agreed, explaining that they conduct these activities only when the lesson topic is appropriate and sufficient time is available. These tasks typically require more complex facilitation, time investment, and higher levels of pedagogical competence, and the relatively lower means suggest structural and instructional constraints that limit their consistent implementation. Literature indicates that teachers often underutilize complex performance tasks due to insufficient training, limited resources, and competing curricular demands (Zaabalawi & Zaabalawi, 2024; Ma'rufah et al., 2024). From a constructivist perspective, collaborative and interactive strategies are critical in fostering dialogic learning and critical discourse, and their moderate use suggests that while teachers recognize their pedagogical value, practical limitations hinder their sustained integration (Wood, 2025; Lin et al., 2025).

The observed pattern reflects a selective adoption of performance tasks, wherein teachers favor activities that are relatively easier to administer and assess while more complex, process-oriented tasks are less frequently utilized. This uneven implementation underscores a misalignment between policy-driven advocacy for authentic assessment and actual classroom practice. As highlighted in the literature, the successful integration of performance-based assessment depends not only on the availability of tools but also on teachers' assessment literacy and institutional support systems (Sudimantara et al., 2025; Alghammas, 2020). From a theoretical standpoint, the findings suggest partial adherence to constructivist and authentic assessment paradigms, where learning is expected to be active, contextualized, and performance-driven. The predominance of "Sometimes" responses indicates that such paradigms are not yet fully institutionalized in the teaching of Araling Panlipunan 7, and that instructional practices appear to oscillate between traditional and progressive approaches, reflecting a condition of pedagogical hybridity.

With respect to the moderating variable of sex, both male (M = 3.26, SD = 0.594) and female (M = 3.39, SD = 0.489) respondents rated the extent of use of performance tasks within the "Sometimes" range. While female students recorded a slightly higher mean score, the

Kruskal-Wallis test result ($\chi^2 = 2.91$, $df = 1$, $p > 0.088$) indicates no statistically significant difference in perceptions based on sex, as the p-value exceeds the 0.05 threshold. The null hypothesis is therefore accepted. Qualitative insights gathered during interviews explain the slight difference: female respondents shared that they found these activities more meaningful and helpful in understanding lessons and appreciated the opportunities for creativity and real-life application, while male respondents tended to view the tasks as similar to other school requirements without paying close attention to their frequency. Nonetheless, the absence of a statistically significant difference reinforces the descriptive findings and suggests that performance task utilization is experienced uniformly regardless of gender. This convergence reflects the inherently inclusive nature of authentic assessment, which emphasizes multimodal engagement rather than gendered learning preferences. The finding aligns with literature asserting that well-designed performance tasks minimize demographic bias and promote equitable learning environments (Zaabalawi & Zaabalawi, 2024; Deane, 2023; Koh & Fu, 2025; Rehman et al., 2023).

With respect to district affiliation, slight variation in means was observed across the three districts: Lubuagan ($M = 3.44$, $SD = 0.258$), Tanudan ($M = 3.34$, $SD = 0.704$), and Pasil ($M = 3.14$, $SD = 0.428$). While Tanudan and Pasil fall within the "Sometimes" range, Lubuagan approaches the lower boundary of "Often," suggesting relatively stronger implementation in that district. Student interviews shed further light on these differences. Learners from Lubuagan reported that their teachers consistently integrate performance tasks in almost every lesson as part of daily discussions, connecting tasks to local traditions and community experiences to make activities more meaningful. Students from Tanudan noted that their teachers use performance tasks regularly and find them helpful but deploy them more strategically rather than in every discussion. Students from Pasil shared that while tasks are always relevant and enjoyable, they are implemented less frequently compared to the other districts, as teachers prioritize quality and depth over frequency. The Kruskal-Wallis test result ($\chi^2 = 17.5$, $df = 2$, $p < 0.001$) indicates a statistically significant difference across districts, as the p-value falls well below the 0.05 threshold, leading to the rejection of the null hypothesis. This finding confirms that location-based factors influence the extent of performance task utilization and underscores the role of contextual variability in shaping instructional practices, particularly in decentralized education systems. The significant disparity suggests that while performance-based assessment is conceptually endorsed across the division, its implementation is contingent upon local capacities including teacher training and resource access. Literature highlights that authentic assessment practices are highly sensitive to contextual conditions, requiring sustained institutional support to ensure consistency across settings (Lin et al., 2025; Alghammas, 2020; Sudimantara et al., 2025; Wood, 2025).

Level of effectiveness of performance tasks in Araling Panlipunan 7

The second objective of the study sought to determine the level of effectiveness of performance tasks in teaching Araling Panlipunan 7 as perceived by the student respondents. The overall average weighted mean of 3.54, falling within the interval of 3.41 to 4.20, yields a descriptive interpretation of "Often" or "Much Effective," indicating that performance tasks generally enhance learning in Araling Panlipunan 7. This finding affirms the broad pedagogical value of performance-based assessment strategies in developing student understanding, engagement, and conceptual competence within the subject.

Among the 20 individual indicators assessed, Suri-Larawan obtained the highest mean of 4.03 ($SD = 1.10$), followed by Loop a Word ($M = 3.82$, $SD = 1.22$) and Slogan Making ($M = 3.80$, $SD = 1.21$), all falling within the "Often" or "Much Effective" range. Additional tasks

rated as "Often" include Kinig-Kanta ($M = 3.73$, $SD = 1.16$), Hula Bandila ($M = 3.70$, $SD = 1.23$), Hula Salita ($M = 3.65$, $SD = 1.28$), Guess the Country or Place ($M = 3.64$, $SD = 1.22$), MaPanuri ($M = 3.63$, $SD = 1.26$), Sanaysay ($M = 3.60$, $SD = 1.35$), Suri-Simbolo ($M = 3.60$, $SD = 1.26$), Ikuwento, Wariin, Hanapin, Lagumin or K-W-H-L Chart ($M = 3.57$, $SD = 1.24$), Founders Keeper ($M = 3.55$, $SD = 1.26$), MAPA-Hanap ($M = 3.51$, $SD = 1.30$), Debate ($M = 3.46$, $SD = 1.32$), and Emoticon Decipher ($M = 3.43$, $SD = 1.37$). Students explained that Suri-Larawan was their favorite because they could easily understand the lesson by examining images and share their own ideas without needing lengthy explanations. For Loop a Word and Slogan Making, they described these tasks as fun and creative, noting that making slogans or finding words related to the topic felt like a game or art activity rather than schoolwork, allowing them to connect prior knowledge to short, memorable expressions that aided retention. These findings align with authentic assessment theory, which posits that performance-based tasks foster meaningful learning by engaging students in real-world, cognitively demanding activities (Koh & Fu, 2025; Rehman et al., 2023).

Several tasks fell within the "Sometimes" or "Moderately Effective" range. Panayam o Talkshow obtained a mean of 3.07 ($SD = 1.32$), Suri-Balita a mean of 3.20 ($SD = 1.31$), Gallery Walk a mean of 3.37 ($SD = 1.27$), Freedom Wall a mean of 3.37 ($SD = 1.33$), and Imperialism Profile a mean of 3.00 ($SD = 1.45$). Students explained that Panayam o Talkshow caused nervousness and anxiety because it required speaking or performing in front of the class, formulating interview questions, and discussing topics they did not yet fully understand. For Imperialism Profile, learners noted that the task required extensive research, complex organization of information about countries, leaders, and events, and that unclear instructions often led them to complete the activity merely for compliance rather than for genuine understanding. Regarding Gallery Walk, students shared that while examining others' work was acceptable, insufficient time and crowded conditions prevented deep reading or meaningful discussion, often resulting in superficial engagement where they simply looked and moved on rather than exchanging substantive ideas. These tasks typically require advanced facilitation skills, scaffolding, and classroom management, which may not be uniformly developed across teachers, and this uneven effectiveness reflects implementation gaps rather than conceptual weaknesses of performance-based assessment. The success of such tasks depends heavily on teacher competence, feedback mechanisms, and instructional design quality (Zaabalawi & Zaabalawi, 2024; Ma'rufah et al., 2024; Wood, 2025; Lin et al., 2025).

Within the Philippine context, these findings reflect a partial institutionalization of authentic assessment, where performance tasks are broadly recognized as effective but are not yet optimally executed across all instructional settings. Structural constraints such as limited training and contextual disparities likely mediate their effectiveness, and sustained professional development and contextualized support systems are critical to maximizing the pedagogical potential of performance-based assessments (Sudimantara et al., 2025; Alghammas, 2020).

With respect to the moderating variable of sex, both male ($M = 3.50$, $SD = 0.526$) and female ($M = 3.58$, $SD = 0.463$) respondents perceived performance tasks as "Much Effective," with a total average weighted mean of 3.54. The negligible difference in means suggests that effectiveness is consistently experienced across gender groups. During interviews, male learners explained that they appreciated these tasks because they were not limited to writing or reading but involved making, drawing, acting, or working with their hands, which helped them understand lessons more concretely. Female respondents similarly valued the tasks for enabling creative expression, collaborative work, and varied forms of demonstrating knowledge, noting that this made them feel that their ideas mattered even when they were not strong test-takers. Both groups agreed that while females appreciated the creative dimension slightly more and

males preferred the hands-on components, they were fundamentally aligned in finding the tasks effective for learning. The Kruskal-Wallis test result ($\chi^2 = 3.35$, $df = 1$, $p > 0.067$) indicates no statistically significant difference in perceived effectiveness across sex, as the p-value exceeds the 0.05 threshold. The null hypothesis is therefore accepted. This finding strengthens the validity of performance-based assessment as a gender-neutral pedagogical approach and supports the universality of constructivist learning processes wherein knowledge construction occurs through shared engagement rather than demographic distinctions. Well-designed performance tasks ensure equitable cognitive access and minimize bias (Zaabalawi & Zaabalawi, 2024; Deane, 2023; Koh & Fu, 2025; Rehman et al., 2023).

With respect to district affiliation, notable variation was observed in the perceived effectiveness of performance tasks: Tanudan obtained a mean of 3.74 ($SD = 0.605$), described as "Much Effective"; Lubuagan obtained a mean of 3.41 ($SD = 0.300$), at the lower boundary of "Much Effective"; and Pasil obtained a mean of 3.29 ($SD = 0.238$), falling within the "Moderately Effective" range. Student interviews illuminated the reasons for these differences. Learners from Tanudan, which recorded the highest rating, explained that their teachers clearly explained activities, provided examples, guided students throughout each task, connected activities to recently discussed content, ensured adequate materials and time, and facilitated group work that made learning enjoyable and effective. Students from Lubuagan acknowledged that activities were still very helpful and made learning easier, but noted that instructions were sometimes unclear or time was insufficient, suggesting room for improvement in delivery. Learners from Pasil, who gave the lowest rating, reported that while tasks were acceptable, they sometimes lacked necessary materials, felt disconnected from their daily lives, were overly complex, and were not always accompanied by sufficient teacher guidance, leading students to complete tasks for compliance rather than genuine understanding. Higher effectiveness in Tanudan likely reflects stronger instructional support and teacher capacity, while lower ratings in Pasil suggest resource and training constraints. The Kruskal-Wallis test result ($\chi^2 = 50.1$, $df = 2$, $p < 0.001$) reveals a statistically significant difference in effectiveness across districts, as the p-value falls well below the 0.05 threshold, leading to the rejection of the null hypothesis. This finding highlights systemic inequities in the implementation of performance-based assessment within decentralized educational contexts and affirms that without sustained institutional investment, the impact of performance tasks remains uneven across settings (Lin et al., 2025; Alghammas, 2020; Sudimantara et al., 2025; Wood, 2025).

Challenges encountered by teachers in the use of performance tasks in Araling Panlipunan 7

The third objective of the study sought to identify the challenges encountered by teachers in implementing performance tasks in Araling Panlipunan 7. The overall average weighted mean of 2.47, falling within the interval of 1.81 to 2.60, yields a descriptive interpretation of "Slightly Serious," indicating that while challenges exist, they are not perceived as severely constraining the implementation of performance tasks. When asked about this general rating, teachers explained that although they face difficulties, they have learned to adjust and find ways to make instruction work, and that their belief in the importance of performance tasks motivates them to adapt their methods or simplify activities to fit their circumstances.

However, this aggregate mean masks critical variations across individual indicators. Limited skills to conduct research ($M = 3.07$, $SD = 0.730$) and limited opportunities for seminars or training to enhance teachers' skills ($M = 3.07$, $SD = 0.997$) both fall within the "Moderately Serious" range of 2.61 to 3.40, identifying teacher capacity as the most salient

constraint in implementation. Teachers highlighted these as their most significant concerns, explaining that many performance tasks require them to design research-based activities and guide students in gathering and analyzing information, but that they were not fully trained to do so effectively. They also noted that seminars, workshops, and training specifically focused on performance-based assessment are very limited, and that most professional development opportunities they receive are general in nature and insufficient for developing the specific competencies required to plan, facilitate, and evaluate these kinds of activities. This finding suggests that the primary challenge is not structural scarcity per se, but insufficient professional development to operationalize authentic assessment effectively. Literature emphasizes that teacher expertise is the primary determinant of successful performance-based assessment implementation (Koh & Fu, 2025; Sudimantara et al., 2025).

In contrast, traditionally cited barriers such as lack of instructional materials ($M = 1.86$, $SD = 0.535$) and lack of school facilities ($M = 2.00$, $SD = 0.555$) were rated as "Slightly Serious," indicating that material and infrastructural conditions are relatively adequate in the study context. Limited ability to meet the minimum learning competency similarly obtained a mean of 2.00 ($SD = 0.555$) and was rated as "Slightly Serious." Teachers explained that materials, facilities, and access to technology are no longer major issues, as all participating schools are equipped with computers, projectors, and stable internet connections that are freely accessible for lessons, presentations, and research activities. They further noted that for hands-on activities, they rely on simple and readily available resources such as cartolina, manila paper, and old newspapers, and that classrooms are functional enough to accommodate group work, rearrangement, and outdoor activities when needed. This shift in the analytical focus toward pedagogical competence rather than resource deficiency aligns with literature emphasizing that teacher expertise is the primary determinant of successful performance-based assessment implementation (Koh & Fu, 2025; Sudimantara et al., 2025). Limited knowledge on the development of instructional materials obtained a mean of 2.43 ($SD = 0.514$), interpreted as "Slightly Serious," further reinforcing the primacy of capacity-related over resource-related constraints.

The moderate concern regarding co-curricular and extra-curricular workload ($M = 2.86$, $SD = 0.663$), falling within the "Moderately Serious" range, reflects systemic pressures that dilute instructional focus and limit teachers' capacity to design and implement complex, process-oriented tasks. Teachers elaborated that their numerous non-teaching responsibilities, including advisories, club activities, meetings, reports, and community work, consume substantial time and energy, often leading them to rush or simplify performance tasks rather than executing them with the depth and detail they would prefer. Performance assessments by nature require sustained planning, scaffolding, and feedback cycles, and these elements are difficult to maintain under heavy workload conditions. Prior studies confirm that time constraints and competing responsibilities significantly hinder the adoption of authentic assessment practices (Ma'rufah et al., 2024; Alghammas, 2020).

From a constructivist perspective, these findings reveal a misalignment between pedagogical ideals and implementation realities. While performance tasks are theoretically positioned to enhance higher-order thinking and contextual learning, their effectiveness is contingent upon teacher readiness and institutional support. The relatively low overall challenge mean of 2.47 may also reflect a normalization of constraints, wherein teachers adapt to limitations rather than perceiving them as critical barriers. Teachers themselves acknowledged that over time they have grown accustomed to these limitations, no longer viewing them as major obstacles but rather as conditions of their daily work that they manage through adaptation. In the Philippine context, particularly in district-based settings, these

findings underscore the need for targeted professional development programs focused on assessment literacy, instructional design, and research competence. Literature consistently highlights that without sustained capacity-building initiatives, the transformative potential of performance-based assessment remains underutilized (Rehman et al., 2023; Lin et al., 2025; Wood, 2025).

Taken together, the findings of this study present a coherent and empirically grounded picture of how performance tasks are utilized, perceived, and constrained in the teaching of Araling Panlipunan 7 across the districts of Tanudan, Pasil, and Lubuagan in the Kalinga Division. The overall average weighted mean of 3.32 on the extent of use confirms a moderate but inconsistent integration of performance-based strategies, while the overall average weighted mean of 3.54 on effectiveness affirms their general pedagogical value as perceived by students. The overall average weighted mean of 2.47 on challenges indicates that teacher capacity, particularly limited research skills and insufficient professional development opportunities, constitutes the most pressing concern, more so than material or infrastructural constraints. Statistically significant differences across districts were confirmed for both extent of use ($\chi^2 = 17.5$, $df = 2$, $p < 0.001$) and effectiveness ($\chi^2 = 50.1$, $df = 2$, $p < 0.001$), underscoring the critical role of contextual and institutional factors in shaping the quality of performance-based assessment implementation. In contrast, no significant differences were found across sex for either extent of use ($\chi^2 = 2.91$, $df = 1$, $p > 0.088$) or effectiveness ($\chi^2 = 3.35$, $df = 1$, $p > 0.067$), affirming the inclusive and gender-neutral character of authentic assessment approaches. These results collectively reinforce the theoretical foundations of constructivist and authentic assessment frameworks while simultaneously exposing the structural and professional development gaps that limit their full realization in Philippine district contexts. The subsequent chapter will present the conclusions drawn from these findings and the recommendations proposed to strengthen the implementation of performance-based assessment in Araling Panlipunan 7 across the Kalinga Division.

CONCLUSION

This study examined the utilization, instructional effectiveness, and implementation challenges of performance tasks in teaching Araling Panlipunan 7 within selected districts in the Kalinga Division of the Philippines. Drawing from the findings presented in the preceding chapter, this concluding discussion synthesizes the key results across the three research objectives, articulates the conclusions derived from the data, and presents recommendations aimed at strengthening the implementation of performance-based assessment in the study context.

With respect to the extent of use, the overall average weighted mean of 3.32, falling within the interval of 2.61 to 3.40 and interpreted as "Sometimes," confirms that performance tasks are moderately but inconsistently integrated into classroom instruction in Araling Panlipunan 7. This finding reflects a transitional stage in assessment practice, wherein teachers have begun incorporating performance-based strategies but have not yet achieved systematic or sustained adoption across all task types. A closer examination of individual indicators reveals notable variability. Tasks such as Slogan Making ($M = 3.86$), Guess the Country or Place ($M = 3.76$), Sanaysay ($M = 3.75$), and Suri-Larawan ($M = 3.72$) fall within the "Often" range of 3.41 to 4.20, indicating relatively strong adoption of cognitively engaging and context-driven activities. In contrast, tasks such as Freedom Wall ($M = 2.76$), Panayam o Talkshow ($M = 2.74$), Suri-Balita ($M = 2.72$), and Imperialism Profile ($M = 2.80$) remain within the lower "Sometimes" range, approaching "Rarely" utilization, reflecting the heavier facilitation demands and time investment these activities require. Collaborative and interactive strategies

such as Debate (M = 3.10) and Gallery Walk (M = 3.02) similarly fall within the "Sometimes" category, highlighting a gap in the implementation of socially mediated learning approaches that are essential to developing critical discourse and disciplinary reasoning. The overall pattern indicates a selective adoption of performance tasks, wherein teachers favor activities that are relatively straightforward to administer and assess while more complex, process-oriented tasks remain underutilized. This uneven implementation underscores a misalignment between policy-driven advocacy for authentic assessment and actual classroom practice, and points to the continuing need for capacity-building and instructional support to move teachers from transitional toward fully institutionalized use of performance-based strategies. It is also noted that no significant differences were found in respondents' perceptions on the extent of use of performance tasks according to the moderating variables of sex, district affiliation, and access to technology, affirming that the overall pattern of moderate and inconsistent utilization is broadly shared across demographic groups in the study context.

With respect to the level of effectiveness, the overall average weighted mean of 3.54, interpreted as "Much Effective" within the range of 3.41 to 4.20, affirms that performance tasks are generally perceived by students as valuable tools for enhancing learning in Araling Panlipunan 7. High-performing tasks such as Suri-Larawan (M = 4.03), Loop a Word (M = 3.82), and Slogan Making (M = 3.80) demonstrate strong effectiveness in promoting higher-order thinking, creativity, and contextual understanding, reflecting their alignment with authentic assessment principles that situate learning in meaningful and engaging contexts. However, certain tasks including Panayam o Talkshow (M = 3.07) and Imperialism Profile (M = 3.00) fall within the "Moderately Effective" range of 2.61 to 3.40, indicating inconsistent instructional impact attributable to the advanced facilitation demands, complex scaffolding requirements, and higher levels of teacher pedagogical competence these activities necessitate. Collaborative strategies such as Gallery Walk (M = 3.37) similarly reveal only moderate effectiveness, pointing to an underutilization of social constructivist learning processes that are critical to developing critical thinking and disciplinary engagement. These findings collectively reflect a partial institutionalization of authentic assessment in the Philippine district context, wherein performance tasks are broadly recognized as effective but are not yet optimally executed across all instructional settings. The evidence suggests that the challenge lies not in the conceptual value of performance-based assessment itself, but in the consistency and quality of its implementation, which is mediated by teacher competence, feedback mechanisms, and the availability of sustained professional support. Consistent with the findings on extent of use, no significant differences were found in respondents' perceptions on the effectiveness of performance tasks according to the moderating variables of sex, district affiliation, and access to technology at the descriptive level, though statistically significant differences across district affiliation were confirmed through the Kruskal-Wallis test, further underscoring the role of localized conditions in shaping instructional outcomes.

With respect to the challenges encountered by teachers, the overall average weighted mean of 2.47, interpreted as "Slightly Serious" within the range of 1.81 to 2.60, indicates that while challenges are present, they are not perceived as severely constraining the overall implementation of performance tasks. However, this aggregate mean masks critical variations that warrant attention. Limited skills to conduct research (M = 3.07) and limited opportunities for seminars or training to enhance teachers' skills (M = 3.07) both fall within the "Moderately Serious" range, identifying teacher capacity as the most salient constraint on effective implementation. These findings affirm that the primary challenge is not structural scarcity but rather the insufficiency of professional development opportunities specifically oriented toward authentic assessment design, facilitation, and evaluation. In contrast, traditionally cited barriers

such as lack of instructional materials ($M = 1.86$) and lack of school facilities ($M = 2.00$) were rated as "Slightly Serious," confirming that material and infrastructural conditions are relatively adequate in the study context and that access to technology, which was uniformly reported across all 287 student respondents, no longer constitutes a meaningful barrier to performance-based instruction. The moderate concern regarding co-curricular and extra-curricular workload ($M = 2.86$), falling within the "Moderately Serious" range, further reflects systemic pressures that dilute instructional focus and limit teachers' capacity to design and sustain complex, process-oriented assessment tasks. Taken together, these findings point to the primacy of professional development, assessment literacy, instructional design competence, and research skills as the critical areas requiring targeted intervention in Philippine district-based educational settings.

Drawing from these findings, the study concludes that performance-based assessment has begun to take meaningful root in the teaching of Araling Panlipunan 7 across the districts of Tanudan, Pasil, and Lubuagan, but its implementation remains uneven across classrooms and task types. Teaching practices continue to reflect a hybrid of traditional and progressive assessment approaches, indicating that the transition toward fully institutionalized authentic assessment is still in progress. The data affirm the inherent pedagogical value of performance tasks in developing higher-order thinking skills, creativity, and deeper conceptual understanding among learners, while simultaneously revealing that this value is not yet being realized consistently or equitably across the division. To address these gaps, it is recommended that the Department of Education and school administrators in the Kalinga Division prioritize the design and delivery of targeted professional development programs focused specifically on performance-based assessment literacy, task design, research-integrated instruction, and classroom facilitation strategies. Schools and district offices are likewise encouraged to create structured opportunities for collegial collaboration among Araling Panlipunan teachers, enabling the sharing of effective practices, peer mentoring, and collaborative development of assessment materials that are contextually relevant and pedagogically sound. Furthermore, workload management policies should be reviewed to ensure that teachers are afforded sufficient time and institutional support to plan, implement, and provide meaningful feedback on performance tasks without being overwhelmed by competing administrative and co-curricular demands. Ultimately, enhancing teacher competence and expanding access to sustained professional development opportunities are essential preconditions for maximizing the transformative potential of performance-based assessment and for advancing student learning outcomes in Araling Panlipunan 7 in alignment with the goals of quality, learner-centered education as articulated in the DepEd MATATAG K-10 Curriculum and the broader commitments of SDG 4 on Quality Education.

REFERENCES

- Alghammas, A. (2020). Online Language Assessment During The COVID-19 Pandemic: University Faculty Members' Perceptions and Practices. *Asian EFL Journal*, 27(44), 169–195.
- Avila-Bellen, E., Baylon Jr., I., Jaravata, J., & Bersabe, N. (2019). Increase The Performance Level of Grade 7 Students in Araling Panlipunan Through Enhanced Assessment Process. *Ascendens Asia Journal of Multidisciplinary Research Abstracts*, 3(2E). <https://doi.org/10.XXXX/AAJMRA.3.2E.6478>

Bayrante, M. F. (2021, August 4). Online Game-Based Strategy in Teaching Araling Panlipunan And the Students' Academic Performance. Zenodo. <https://doi.org/10.5281/zenodo.5157784>

Bunglo, A. A., & de Guzman, M. F. D. (2023). Effectiveness And Usefulness Of Student Worksheets as Secondary Social Studies Teaching Resource. *Asian Journal of Social Science and Management Technology*, 5(3), 87–100.

Burayag, A. L. (2024). Managing Competencies of The Araling Panlipunan 7 Through The Use of Mobile-Assisted Learning Materials. *Psychology and Education: A Multidisciplinary Journal*, 25(4), 565–569. <https://doi.org/10.5281/zenodo.13785717>

Cañete, J. B. (2024). School Performance in Araling Panlipunan Among Grade 7 Students in San Miguel 1 District, Surigao Del Sur, Philippines. *VMCA International Journal of Education and Teaching Across Sciences*, 3(1), 98–113. <https://doi.org/10.62718/vmca.pr-ijetas.3.1.SC-0924-010>

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approach* (5th Ed.). SAGE Publications.

Darling-Hammond, L., & Adamson, F. (2014). *Beyond The Bubble Test: How Performance Assessments Support 21st Century Learning*.

David, R. M. T. (2024). Teaching Araling Panlipunan: A Study of Challenges and Coping Mechanisms Among Non-Social Studies Major Educators. *International Journal of Multidisciplinary: Applied Business and Education Research*, 5(8), 3339–3348. <http://dx.doi.org/10.11594/ijmaber.05.08.29>

Deane, P. D. (2023). *Achieving Equity in School Writing: Causes and Cures for Opportunity and Achievement Gaps in a Key Twenty-First Century Skill*. Routledge.

Department of Education. (2019). DepEd Order No. 21, s. 2019: Policy Guidelines on the K to 12 Basic Education Program. <https://www.deped.gov.ph/2019/05/30/do-21-s-2019-policy-guidelines-on-the-k-to-12-basic-education-program/>

Department of Education. (2024). MATATAG Curriculum Guide: Araling Panlipunan Grades 4 and 7. <https://www.deped.gov.ph/wp-content/uploads/MATATAG-Araling-Panlipunan-CG-Grade-4-and-7.pdf>

Galvez, M. V. (2024, October). Instructional Efficacy of Araling Panlipunan 7 Teachers: Basis for Intervention Strategies [Issue 10]. *International Journal of Novel Research and Development*, 9(10), Article IJNRD2410199. ISSN 2456-4184.

Koh, K., & Fu, R. (2025). Authentic and Alternative language assessment: Design and Implementation. In *Routledge Handbook of Language Program Development and Administration* (pp. 161–173). Routledge.

Lin, C.-H., Zhou, K., Li, L., & Sun, L. (2025). Integrating Generative AI into Digital Multimodal Composition: A Study of Multicultural Second-Language Classrooms. *Computers and Composition*, 75, 102895. <https://doi.org/10.1016/j.compcom.2024.102895>

Ma'rufah, D. W., Mualim, M., & Hilman, F. A. (2024). Online Peer Assessment in Paragraph Writing: Its Implementation and Students' Constraints. *JALT CALL Journal*, 20(1), 1–15.

Padayao, E. N. (2026). Effectiveness of Learning Activity Sheets in Teaching Araling Panlipunan for Grade 9 Students in Public Secondary Schools. *International Journal of Sustainability and Advanced Integrated Research*, 2(2), 469.[doi.org](https://doi.org/10.1016/j.compcom.2024.102895)

Rehman, N., Zhang, W., Mahmood, A., Fareed, M. Z., & Batool, S. (2023). Fostering Twenty-First Century Skills Among Primary School Students Through Math Project-Based Learning. *Humanities and Social Sciences Communications*, 10(1), 424. <https://doi.org/10.1057/s41599-023-01863-8>

Siedlecki, S. L. (2020). Understanding Descriptive Research Designs and Methods. *Clinical Nurse Specialist*, 34(1), 8-12. <https://doi.org/10.1097/NUR.0000000000000493>

Singca, M. D. (2022). Learning Resources and Learning Competencies In Teaching Araling Panlipunan Towards Student's Outcomes. *International Journal of Research Publications*, 104(1). <https://doi.org/10.47119/ijrp1001041720223530>

Sokhanvar, Z., Salehi, K., & Sokhanvar, F. (2021). Advantages of Authentic Assessment for Improving the Learning Experience and Employability Skills of Higher Education Students: A Systematic Literature Review. *Studies in Educational Evaluation*, 70, Article 101030. <https://doi.org/10.1016/j.stueduc.2021.101030>

Sudimantara, L. B., Baehaqi, L., Lian, A., & Lian, A.-P. (2025). *Multisensory CALL for Under-Resourced Universities and Schools in Indonesia*. Springer.

Sumaylo, J. R., & Espacio, N. B. (2025). DEPED Order (Do) No. 010, S. 2024: Unveiling the Lived Experiences of Araling Panlipunan Teachers in the Implementation of Matatag Curriculum. *International Journal of Research in Innovative Social Sciences and Studies*, 9(4), 5783 5804. <https://doi.org/10.47772/IJRISS.2025.90400412>

Wood, J. (2025). Creativity and Historical Inquiry in Digital Spaces: Perspectives from the Practice of Online Making. *History*, 110(393), 702–712. <https://doi.org/10.1111/1468-229X.13345>

Zaabalawi, R. S., & Zaabalawi, J. (2024). Portfolios Versus Exams: A Study To Gauge The Better Student Assessment Tool. *Language Testing in Asia*, 14(1), 28. <https://doi.org/10.1186/s40468-024-00252-0>

Zubieto, A. A. (2021–2022). Teaching Social Studies in the Elementary Grades [Course Module]. Southern Luzon State University <https://www.scribd.com/document/767056604/aa-ele03-module-1st-sem-2021-2022-2>