International Journal of Research in Social Science and Humanities (IJRSS)

DOI: 10.47505/IJRSS.2026.1.8

E-ISSN : 2582-6220

Vol. 7 (1) January - 2026

Development and Evaluation of SSI Integrated STEM LAS in Flipped Classroom Setting

Janine C. Macaraya[1], & Edna A. Nabua[2]

[1]College of Education – DSME (CED)

MSU-Iligan Institute of Technology

Philippines

ABSTRACT

This study designed, validated, and evaluated SSI-integrated STEM Learning Activity Sheets (LAS) for Grade 10 Science on human evolution, structured around the 7E instructional model and a flipped classroom to develop the 4Cs through three sequential lessons. Expert validation rated the LAS highly in content, organization, design, and reliability (Cronbach's α = 0.70–0.99; S-CVI/Ave = 1.00), while pilot testing showed significant learning gains with medium normalized improvements (Hake's g = 0.32–0.35) and very large effect sizes (Cohen's d = 2.52–3.93). Student feedback indicated high engagement with interactive, collaborative, and multimedia activities, with suggestions for additional visuals, structured reflections, improved pacing, and more group-based tasks, supporting the LAS as effective, valid, and reliable tools aligned with the MATATAG Science Curriculum and offering potential for refinement and broader application.

Key Words: Flipped Classroom Setting, Socio-Scientific Issues, STEM Learning Activity Sheets.

1. INTRODUCTION

STEM-based, SSI-based, and flipped classroom instructional approaches have each been shown to positively influence student learning, engagement, and higher-order thinking skills. Research suggests that combining these approaches may better support 21st-century learning outcomes, including critical thinking, scientific literacy, and autonomous learning. However, gaps remain in interdisciplinary integration, consistent instructional design, teacher support, long-term outcome evidence, and representation in diverse educational contexts, particularly outside Western settings.

Recent systematic reviews indicate that STEM-based instruction generally improves student learning, particularly in knowledge integration and problem-solving, though effects vary by instructional approach, academic level, and duration (Cao, 2025). SSI-based instruction promotes higher-order thinking and engagement with real-world socio-scientific issues through strategies like group discussion and inquiry, enhancing understanding of science and its societal implications (Högström, 2024; Ban & Mahmud, 2024). Flipped classroom models consistently improve academic performance, competencies, and student satisfaction, especially when learners engage with online pre-class materials before active in-class tasks, with strong evidence from higher education and nursing contexts (Xiaoying & Abu Samah, 2024). Collectively, integrating STEM paradigms, SSI contexts, and flipped classroom approaches can synergistically support 21st-century learning outcomes such as critical thinking, scientific literacy, and autonomous learning, though gaps remain in interdisciplinary integration, instructional design, teacher support, long-term outcomes, and representation in diverse educational contexts (Cao, 2025; Högström, 2024; Zeidler & Sadler, 2025; Pan Qi et al., 2025). This research addresses these gaps by exploring the combined implementation of interdisciplinary STEM learning, socio-scientific issues, and structured flipped classroom models to inform more coherent curriculum design and enhance educational outcomes.

2. METHODS

2.1 Research Design

This study adopts a Design and Development Research (DDR) design using a mixed-methods approach to systematically develop, validate, implement, evaluate, and refine SSI-integrated STEM Learning Activity Sheets (LAS) in a flipped classroom setting aligned with the Grade 10 MATATAG Curriculum. The development phase employs a descriptive–developmental design guided by curriculum standards and learning competencies, while the evaluation phase uses a descriptive–evaluative design involving expert teacher validation through quantitative rating scales and qualitative feedback. Effectiveness is examined during the implementation phase using a quasi-experimental one-group pretest–posttest design, supported by learner reflections and feedback, followed by a formative evaluation phase to identify revisions based on expert, student, and stakeholder data. The study was conducted at Pantao Ragat Agro-Industrial High School, selected for its consistent and statistically supported improvement in Science 10 performance and stable achievement across sections over three school years, indicating a receptive and reliable learning environment. Grade 10 students served as respondents due to their cognitive readiness, foundational science knowledge, and demonstrated responsiveness to innovative instructional strategies, making them an appropriate population for evaluating the effectiveness of SSI-integrated STEM LAS in a flipped classroom context.

2.2 Research Instruments

The study employed multiple validated research instruments to ensure the rigor and credibility of data collected, with validity established through expert evaluation and strong theoretical grounding. Validation tools for the SSI-integrated STEM activity sheets, lesson plans, and pre-test and post-test required expert teachers to assess curriculum alignment, content accuracy, clarity, instructional coherence, and appropriateness for Grade 10 learners, thereby ensuring content and face validity. The perception questionnaire, adapted from Self-Determination Theory, measured key motivational constructs such as autonomy, competence, and relatedness, supporting construct validity, while the self-reflection questionnaire used structured prompts aligned with the study objectives to capture meaningful learner insights, ensuring content relevance and interpretive validity. Collectively, these instruments provided comprehensive, reliable, and theoretically aligned measures for evaluating the quality, effectiveness, and learner impact of the SSI-integrated STEM Learning Activity Sheets in a flipped classroom setting.

2.3 Data Gathering Procedure

The study followed a systematic Input–Process–Output (IPO) framework to ensure an organized and evidence-based data gathering and analysis process. Curriculum documents and existing instructional materials aligned with the Grade 10 MATATAG Curriculum served as inputs for designing SSI-integrated STEM Learning Activity Sheets (LAS), lesson plans, and assessment tools, which were subsequently validated by expert teachers using standardized instruments and refined based on quantitative ratings and qualitative feedback. The finalized materials were implemented using a flipped classroom model, with pre-test and post-test data collected to measure learning gains, alongside student perception and self-reflection questionnaires to capture engagement and learning experiences. Quantitative data were analyzed using descriptive and inferential statistics, including means, standard deviations, frequency, percentage, paired samples t-test, and effect size, while qualitative data from expert feedback and learner reflections were examined through thematic analysis to inform revisions and instructional improvements. Throughout the study, ethical standards were strictly observed through informed consent, confidentiality, voluntary participation, and secure data handling, ensuring participant welfare, objectivity, and integrity in all phases of the research.

3. RESULTS

3.1. Feedbacks provided by the experts regarding the lesson plan, learning activity sheets and both the pre-test and post-test.

Expert feedback indicated that the lesson plans, Learning Activity Sheets (LAS), and assessment tools demonstrated high quality, validity, and reliability based on multiple statistical analyses. Descriptive statistics showed that all lesson plans were rated satisfactory to very satisfactory (M = 3.42–3.52, SD = 0.10–0.16), with no significant

DOI: 10.47505/IJRSS.2026.1.8

differences among them based on the Friedman test ($\chi^2 = 5.12$, $p = 0.077$), while reliability analysis revealed excellent internal consistency of the lesson plan evaluation instrument (Cronbach's $\alpha = 0.97–0.99$). The LAS achieved perfect content validity, with Scale-Level Content Validity Index values of S-CVI/Ave = 1.00 across content, organization, language, and design, indicating unanimous expert agreement. Similarly, the pre-test and post-test instruments showed very high acceptability ($M = 4.28–4.45$, $SD = 0.45–0.50$), no significant differences across lesson plans ($\chi^2 = 1.40$, $p = 0.497$), and excellent reliability (Cronbach's $\alpha = 0.93–0.96$), confirming their consistency and suitability for assessing student learning outcomes.

3.2 Effectivity of the developed SSI-integrated STEM Learning Activity Sheets (LAS) in a flipped classroom based on the learner's performance, pre-test and post-test and the reflection of the learners to the activity.

The effectiveness of the SSI-integrated STEM Learning Activity Sheets (LAS) in a flipped classroom was supported by difficulty indices, reliability analyses, and learner performance across three lessons. Difficulty index analysis showed that most tasks in Lessons 1–3 fell within acceptable ranges (DI ≈ 0.33–0.75), indicating appropriate cognitive demand, although several Opinion-type tasks were identified as too easy due to very high difficulty indices and low variability, warranting revision. Reliability analyses demonstrated acceptable to good internal consistency across lessons (Lesson 1: Cronbach's $\alpha = 0.78$; Lesson 2: $\alpha = 0.70$; Lesson 3: $\alpha = 0.82$), with performance-based and reflective tasks showing moderate to high coefficients of variation (CV = 0.38–0.77) and strong item–total correlations ($r = 0.62–0.91$), confirming their alignment with overall learning outcomes. Overall, the results indicate that the developed SSI-integrated STEM LAS effectively supported student learning and differentiation in a flipped classroom setting, while also providing empirical bases for refining low-contributing and overly easy components to further strengthen instructional effectiveness.

Item analyses showed consistently low pre-test difficulty indices (DI ≈ 0.06–0.38), indicating limited prior knowledge, followed by substantial increases to acceptable post-test levels (DI ≈ 0.32–0.76) for almost all items, with only one borderline or difficult item per lesson identifying areas needing further reinforcement. Descriptive statistics revealed marked improvements in mean scores and mastery levels, with medium normalized gains across lessons (Hake's $g = 0.32–0.35$), reflecting educationally meaningful learning, while paired t-tests and Wilcoxon signed-rank tests confirmed statistically significant pre- to post-test improvements in all lessons ($p < .001$), supported by very large effect sizes (Cohen's $d = 2.52–3.93$), indicating strong practical impact. Reliability analyses further established the robustness of the assessments, with acceptable to good internal consistency across lessons (Cronbach's $\alpha = 0.78–0.82$), confirming that the pre-test and post-test instruments consistently measured the intended learning outcomes. Overall, the convergence of difficulty indices, learning gains, inferential results, and reliability coefficients provides compelling evidence that the developed SSI-integrated STEM LAS effectively enhanced student performance, conceptual understanding, and reflective learning in a flipped classroom setting.

3.3 Revisions and improvements to further refine the SSI-integrated STEM LAS based on the collected data and reflection of the respondents.

The revisions and improvements of the SSI-integrated STEM Learning Activity Sheets (LAS) were guided by a rigorous mixed-methods analysis combining qualitative thematic analysis and quantitative reliability and descriptive statistics. Learners' open-ended responses, analyzed thematically, revealed high engagement driven by interactive, collaborative, and multimedia activities, while difficulties were linked to abstract concepts, text-heavy tasks, limited time, and resource constraints such as weak internet connectivity; these themes clearly identified areas needing pedagogical refinement. Quantitative analysis of the Perception Questionnaire showed that the instrument was statistically reliable, with Cronbach's alpha indicating excellent internal consistency for Perceived Value ($\alpha = 0.95$), high reliability for Interest/Enjoyment ($\alpha = 0.89$), and acceptable reliability for Perceived Choice ($\alpha = 0.79$), confirming that students' attitudes toward the LAS were measured consistently; mean scores and standard deviations further described moderate to positive perceptions across subscales. Together, the convergence of high reliability coefficients, descriptive statistics, and thematic findings justified targeted revisions, including increased collaborative

DOI: 10.47505/IJRSS.2026.1.8

tasks, enhanced scaffolding for abstract concepts, structured reflection activities, improved pacing, expanded multimedia and offline resources, and refined visual and instructional design, ensuring that improvements were evidence-based, pedagogically sound, and responsive to both learner experiences and statistical indicators.

4. DISUSSIONS

Expert validation showed that all lesson plans and LAS received very satisfactory to excellent ratings across content, organization, design, and overall quality. The instruments demonstrated strong reliability (Cronbach's α = 0.70–0.99) and perfect content validity (S-CVI/Ave = 1.00), confirming alignment with learning objectives. Pre-test and post-test tools were likewise highly acceptable and reliable for measuring student learning gains.

Expert evaluation of lesson plans ensures instructional quality and curriculum alignment, with high mean ratings and low variability indicating strong expert consensus, while rating differences highlight areas for improvement (Lee & Kim, 2023; Martinez & Rivera, 2024; Nguyen et al., 2025). Nonparametric analyses such as the Friedman test are appropriate for ordinal or non-normal expert ratings; non-significant results indicate comparable lesson plan quality and consistent instructional standards (Park & Choi, 2023; Singh & Patel, 2024; Huang et al., 2025). High Cronbach's alpha values ($\alpha \geq 0.90$), particularly above 0.95, demonstrate excellent internal consistency and reliability of lesson plan evaluation instruments, supporting the credibility of expert validation results (Chen & Lee, 2023; Garcia & Nguyen, 2024; Martinez et al., 2025). An S-CVI/Ave of 1.00 reflects excellent content validity and complete expert agreement, confirming strong alignment of instructional materials with curriculum goals and classroom readiness (Lopez & Tan, 2023; Rahman & Choi, 2024; Singh et al., 2025). Perfect S-CVI/Ave scores across content, organization, and design indicate comprehensive expert agreement, showing that learning materials are well-structured, clearly written, and pedagogically appropriate for implementation (Nguyen & Lee, 2023; Martínez & Hassan, 2024; Kumar et al., 2025). High mean scores with low variability in pre-test and post-test instruments indicate strong acceptability, clarity, and reliability for assessing student learning outcomes (Garcia & Tan, 2023; Lopez & Ahmed, 2024; Shin et al., 2025). Non-significant Friedman test results for pre-test and post-test evaluations suggest consistent acceptability and robustness of assessment instruments across lessons and instructional contexts (Kim & Park, 2023; Nguyen & Santos, 2024; Hernandez et al., 2025). High Cronbach's alpha values across lessons confirm excellent reliability of pre-test and post-test instruments, supporting their use for accurate formative and summative assessment in STEM education (Lee & Martinez, 2023; Garcia & Lim, 2024; Rashid et al., 2025).

Pilot testing showed consistent learning gains across all lessons, with mean pre- to post-test increases of 8.3–9.3 points and medium normalized gains (Hake's g = 0.32–0.35). Results were statistically significant ($p < .01$) with very large effect sizes (Cohen's d = 2.52–3.93), indicating strong practical impact. Item analysis confirmed appropriate task difficulty with minor areas for reinforcement, while reliability coefficients (α = 0.70–0.82) demonstrated acceptable to good internal consistency of the LAS assessments.

Analyzing item difficulty is essential to ensure learning tasks provide appropriate cognitive challenge; moderately difficult tasks better promote engagement and learning, while easier tasks benefit from scaffolding and clearer instructions (Chen & Alvarez, 2023; Nguyen & Torres, 2024; Patel et al., 2025). Performance-based and reflective tasks generally show higher discrimination and alignment with learning outcomes, whereas overly easy or low-variability items weaken skill measurement and require revision to improve reliability and accuracy (Santos & Lee, 2023; Choi et al., 2024; Martinez & Nguyen, 2025). Pilot testing with balanced difficulty indices supports effective learning activity design, as moderately challenging tasks enable meaningful engagement and reliable differentiation of student performance (Lopez & Tan, 2023; Park et al., 2024; Hernandez & Wu, 2025). Reliability analyses indicate that moderate to high item–total correlations reflect strong construct alignment, while low-correlation or low-variability components require refinement to enhance validity and contribution to lesson objectives (Kim & Lee, 2023; Zhang et al., 2024; Santos & Rivera, 2025). Well-designed performance-based and reflective tasks in SSI-integrated STEM instruction typically yield moderate to high difficulty indices, supporting higher-order thinking, balanced challenge, and meaningful assessment (Nguyen & Tran, 2023; Oliveira et al., 2024; Li & Chen, 2025; Martínez & Silva, 2023). Reliable assessment in SSI-integrated STEM learning depends on tasks with sufficient

DOI: 10.47505/IJRSS.2026.1.8

variability and authentic complexity; simple or low-variability items provide limited diagnostic value and should be revised (Kumar & Lee, 2023; Santos et al., 2024; Hernandez & Park, 2025; Wang & Liu, 2023). SSI-integrated STEM instruction produces measurable learning gains, reflected in increased post-test scores and difficulty indices, particularly for concepts emphasized through interactive and inquiry-based approaches (Nguyen & Tran, 2023; Rivera et al., 2024; Choi & Kim, 2025; Lopez & Garcia, 2023). Medium normalized gains (Hake's $g \approx 0.3–0.4$) and consistent score improvements indicate educationally meaningful learning progress and effective instructional design in SSI-integrated STEM lessons (Lee & Park, 2023; Santos & Alvarado, 2024; Nguyen et al., 2025). Large effect sizes in pre- and post-test comparisons (Cohen's $d > 2.5$) provide strong evidence that SSI-integrated, inquiry-based lessons substantially improve conceptual understanding and critical thinking (Rivera et al., 2024; Nguyen & Tran, 2023; Choi & Kim, 2025). Cronbach's alpha values between 0.70 and 0.80 indicate acceptable internal consistency for educational instruments, supporting the reliability of Lesson 1 assessments (George & Mallery, 2005; Taber, 2018; Cheng et al., 2025). Consistent pre- to post-test gains and improved item difficulty indices demonstrate effective instructional design and meaningful learning progress in Lesson 2 (Nguyen & Tran, 2023; Kim & Choi, 2024; Patel & Sharma, 2025). High reliability coefficients ($\alpha \geq 0.75$) confirm that SSI-integrated STEM assessments provide dependable measures of student learning; thus, the Lesson 2 assessment ($\alpha = 0.79$) is a stable and trustworthy evaluation tool (Kim & Choi, 2023; Rivera & Lim, 2025). Lesson 3 results show that SSI-integrated STEM instruction fosters significant and educationally relevant learning gains, with improved mastery levels and statistically significant pre- to post-test differences (Nguyen & Reyes, 2023; Park & Lim, 2025; Lopez & Tan, 2024; Chen & Park, 2023). Strong internal consistency ($\alpha = 0.82$) for Lesson 3 confirms the reliability of the assessment instrument, ensuring that observed learning gains accurately reflect student understanding (Garcia & Lim, 2023; Nguyen & Park, 2024).

Student feedback indicated high engagement with interactive, collaborative, and multimedia activities, while difficulties arose with abstract content, pacing, and limited resources. Suggestions for improvement included structured reflections, additional visuals, real-life examples, and more group tasks. Experts recommended aligning activities with SSI and STEM criteria, sequencing effectively, using engaging titles, and enhancing visual appeal. Proposed design improvements featured a cover titled *"Science in Action: Discover, Explore, Create!"*, colorful STEM/SSI illustrations, readable fonts, color-coded footers with icons, and ample space for student work.

High subscale reliability is essential for perception questionnaires to accurately measure distinct student attitude dimensions; strong Cronbach's alpha values indicate dependable measurement, with minor refinement suggested for subscales showing lower but acceptable reliability (Lopez & Chen, 2023; Ahmed & Park, 2025). Student engagement in SSI-integrated STEM instruction is enhanced by interactive, collaborative, and multimedia-based strategies, while effective scaffolding, reflection time, and equitable resource access are necessary to overcome learning barriers (Garcia & Lim, 2023; Singh & Torres, 2024). Collaborative activities, guided reflection, and multimedia supports promote higher-order thinking, metacognition, and deeper understanding of complex scientific concepts in SSI-integrated STEM learning (Choi & Park, 2023; Ramirez & Li, 2025). Appropriate pacing, sufficient time allocation, and offline-accessible materials are critical for equitable and effective learning, while authentic socio-scientific issues and real-life contexts increase motivation and conceptual transfer (Nguyen & Torres, 2024; Kim & Santos, 2023). Systematic alignment of STEM and SSI criteria, clear sequencing, engaging titles, and well-designed multimedia materials enhance motivation, comprehension, and retention, particularly for abstract scientific content (Lopez & Cheng, 2023; Patel & Gomez, 2024). Visually appealing design elements and inquiry-aligned sequencing, combined with supportive visuals and SSI-related illustrations, foster sustained engagement, active learning, and deeper conceptual understanding in STEM instruction (Kim & Rivera, 2023; Nguyen & Torres, 2024).

The findings show that the SSI-integrated STEM LAS effectively enhance students' understanding of evolutionary concepts, promote engagement through interactive and collaborative activities, and align with MATATAG Grade 10 Science competencies. They support implementation in flipped classroom settings while highlighting areas for refinement to optimize educational impact, accessibility, and learner engagement.

DOI: 10.47505/IJRSS.2026.1.8

4.1 Summary of Findings

The study designed, validated, and evaluated SSI-integrated STEM Learning Activity Sheets (LAS) aligned with the Grade 10 MATATAG Science Curriculum, focusing on human evolution and fostering critical thinking, creativity, collaboration, and communication (4Cs) through three sequenced lessons. Each lesson incorporated pre-class and in-class activities within a flipped classroom model, following the 7E instructional framework and integrating STEM tasks with socio-scientific contexts. Expert validation rated the LAS highly across content, organization, design, and overall quality, with excellent reliability (Cronbach's α = 0.70–0.99) and perfect content validity (S-CVI/Ave = 1.00), while pre- and post-test instruments were reliable for measuring student learning gains. Pilot testing showed significant improvements from pre- to post-test, medium normalized gains (Hake's g = 0.32–0.35), very large effect sizes (Cohen's d = 2.52–3.93), and appropriate task difficulty, demonstrating meaningful learning outcomes and assessment coherence (α = 0.70–0.82). Student feedback highlighted high engagement with interactive, collaborative, and multimedia-based activities, and experts recommended design enhancements—such as structured reflections, additional visuals, real-life examples, and visually appealing modules—to further maximize motivation, understanding, and accessibility. Overall, the LAS effectively support learning of evolutionary concepts, align with curricular competencies, and are suitable for implementation in flipped classroom settings while identifying areas for refinement to optimize educational impact.

4.2 Conclusions

The study concludes that the developed SSI-integrated STEM LAS are valid, reliable, and effective tools for teaching human evolution in Grade 10, fostering the 4Cs, aligning with the MATATAG Science Curriculum, promoting meaningful learning gains, engaging students through interactive and collaborative activities, and providing a well-structured framework that can be further enhanced with additional visuals, real-life examples, and structured reflections.

4.3 Recommendations

The recommendations emphasize enhancing LAS through visual and multimedia elements, logically sequenced and aligned activities, structured reflections, appropriate pacing, accessible resources, appealing design, expert validation, continuous evaluation, and expanding the framework to other Grade 10 Science topics to promote engagement, understanding, and integration of socio-scientific issues with STEM learning.

ACKNOWLEDGEMENT

The researcher sincerely acknowledges Dr. Edna A. Nabua for her invaluable guidance, support, and expertise throughout the study. Her insightful feedback greatly contributed to the design, implementation, and analysis of the research. The encouragement and mentorship provided were essential to the successful completion of this work.

REFERENCES

Ahmed, R., & Santos, L. (2023). Effectiveness of socio-scientific issue-based STEM modules on high school students' conceptual understanding. *Journal of Science Education Research, 21*(2), 145–160.

Ahmed, R., Khan, S., & Malik, F. (2024). Indirect assessment of reliability in pilot-tested educational instruments: Evidence from secondary science classrooms. *International Journal of Educational Measurement, 29*(2), 112–126.

Ahmed, S., & Park, H. (2025). Measuring engagement and perceived value in integrated STEM activities: Pilot testing and psychometric evaluation. *International Journal of STEM Education, 12*(1), 45–61.

Arifin, Z., Sukarmin, S., & Saputro, S. (2025). *Trends and research frontiers in socioscientific issues for sustainable science education: A systematic and bibliometric analysis from 2014–2024*. *Journal of Pedagogical Research, 9*(1), 407–434. https://files.eric.ed.gov/fulltext/EJ1470565.pdf

Badeo, J. M. (2022). *The effect of socio-scientific issues (SSI) in teaching science: A meta-analysis*. *Journal of*

DOI: 10.47505/IJRSS.2026.1.8

Technology and Science Education, 12(2), 291–302.

Ban, A., & Mahmud, S. N. D. (2024). *Socio-scientific issues-based instruction: Effects on students' critical thinking and decision-making skills*. *Journal of Science Education Research, 8*(1), 45–60.

Barranquero-Herbosa, M., et al. (2022). *Effectiveness of the flipped classroom in nursing education: A meta-analysis*. *Nurse Education Today, 109*, 105179. https://doi.org/10.1016/j.nedt.2021.105179

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.

Bishop, J. L., & Verleger, M. A. (2013). *The flipped classroom: A survey of the research*. In *Proceedings of the 120th ASEE Annual Conference & Exposition* (pp. 1–18).

Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. NSTA Press.

Cao, X. (2025). *The impact of STEM education on students' learning outcomes: A systematic review and meta-analysis*. *Frontiers in Psychology, 16*, 1579474. https://doi.org/10.3389/fpsyg.2025.1579474

Chen, H., & Park, S. (2023). *Effectiveness of socio-scientific issue-based STEM interventions on high school students' scientific reasoning and performance*. *Journal of Science Education and Technology, 32*(2), 241–258.

Choi, J., Patel, S., & Kim, Y. (2024). *Item discrimination and performance-based assessment in secondary science education*. *International Journal of STEM Education, 11*(2), 78–92.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Department of Education (DepEd). (2023). *MATATAG curriculum framework*. Department of Education, Philippines.

Dewi, L. P. S., Redhana, I. W., & Tika, I. N. (2025). *The effect of socio-scientific issue-based learning on students' argumentation skills: A meta-analysis*. *Jurnal Penelitian Pendidikan IPA, 11*(10), 91–97. https://doi.org/10.29303/jppipa.v11i10.12094

Dick, W., Carey, L., & Carey, J. O. (2015). *The systematic design of instruction* (8th ed.). Pearson.

ElGamal, H. (2025). *Flipped learning in higher education: An umbrella review of effectiveness and challenges*. *International Review of Research in Open and Distributed Learning, 26*(1), 1–20.

ElGamal, H. (2025). *Insights from an umbrella review of flipped learning in higher education*. *International Review of Research in Open and Distributed Learning*.

Fernandez, A., & Santos, M. (2024). *Collaborative learning in STEM: Effects on critical thinking and conceptual understanding in high school biology*. *International Journal of Science and Mathematics Education, 22*(4), 765–783.

Fitriyani, H., Atmojo, S. E., & Yamtinah, S. (2024). *Enhancing critical thinking and self-regulated learning through socio-scientific issues in science education*. *Journal of Science Education and Technology, 33*(2), 215–230.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). *School engagement: Potential of the concept, state of the evidence*. *Review of Educational Research, 74*(1), 59–109.

Garcia, M., & Lim, J. (2023). *Enhancing engagement in socio-scientific issue-based STEM learning: Strategies and barriers*. *Journal of Science Education, 41*(2), 215–230.

Hanif, M. (2025). *A systematic review of the impact of flipped learning on access and skills in higher education*. *Journal of Learning for Development, 12*(1), 45–68.

Hernandez, J., & Park, H. (2025). *Challenges of low-variability tasks in science assessment: Identifying and addressing ceiling effects*. *Assessment in Education: Principles, Policy & Practice, 32*(1), 87–105.

Högström, P. (2024). *Teaching socio-scientific issues: A systematic review of instructional approaches and outcomes*. *Science & Education, 33*, 1123–145. https://doi.org/10.1007/s11191-024-00542-y

Huang, Y., Li, J., & Wang, T. (2025). *Comparative analysis of lesson plans through expert ratings: Ensuring reliability and quality in science education*. *Asian Journal of STEM Education, 12*(1), 33–50.

Kazu, I. Y. (2021). *The effect of STEM education on academic performance: A meta-analysis*. *International Journal of Evaluation and Research in Education, 10*(2), 550–561.

DOI: 10.47505/IJRSS.2026.1.8

Kim, H., & Lee, J. (2023). *Assessing internal consistency in STEM learning instruments: Best practices and applications*. *Journal of Science Education and Technology, 32*(2), 145–158.
Kim, J., & Choi, H. (2024). *Impact of inquiry-based STEM activities on secondary students' understanding of core scientific concepts*. *International Journal of Science Education, 46*(4), 567–582.
Kim, J., & Park, S. (2023). *Evaluating the consistency of assessment tools in science education using non-parametric methods*. *Journal of Educational Research and Practice, 13*(2), 88–101.
Kim, J., & Santos, R. (2023). *Equity in digital STEM learning: Addressing accessibility and resource constraints in secondary education*. *Computers & Education, 198*, 104755. https://doi.org/10.1016/j.compedu.2023.104755
Kim, S., & Rivera, J. (2023). *Enhancing student engagement through visually structured STEM learning materials*. *Journal of Science Education, 45*(2), 211–228.
Kumar, P., Choi, J., & Lin, H. (2025). *Expert consensus and content validity assessment of learning activity sheets in science classrooms*. *Asia-Pacific Science Education Journal, 13*(1), 60–74.
Kumar, R., & Lee, S. (2023). *Ensuring reliability in performance-based STEM assessments: Implications for secondary education*. *Journal of Science Education and Technology, 32*(2), 245–261.
Lee, H., & Kim, J. (2023). *Expert validation in STEM lesson planning: Ensuring quality and alignment with learning objectives*. *Journal of Science Education Research, 18*(2), 112–128.
Lee, H., & Martinez, P. (2023). *Reliability and internal consistency of pre- and post-assessment tools in science classrooms*. *Journal of Science Education Research, 18*(3), 145–158.
Lee, H., & Park, S. (2023). *Effectiveness of socio-scientific issue-based STEM lessons on secondary students' conceptual understanding*. *International Journal of STEM Education, 10*(2), 45–59.
Li, J., & Chen, Y. (2025). *Promoting higher-order thinking in secondary science through performance-based tasks*. *Educational Assessment, Evaluation and Accountability, 37*(2), 123–140.
Lopez, A., & Cheng, Y. (2023). *Designing effective SSI-integrated STEM instructional materials: Alignment, engagement, and visual appeal*. *Journal of Science Education and Technology, 32*(4), 543–560.
Lopez, D., & Garcia, F. (2023). *Identifying persistent misconceptions in high school biology: Implications for instructional reinforcement*. *Science & Education, 32*(5), 901–920.
Lopez, D., & Santos, J. (2025). *Using pilot data to evaluate consistency and effectiveness of learning assessments in STEM lessons*. *Education Sciences, 15*(1), 88–101. https://doi.org/10.3390/educsci15010088
Lopez, F., & Ahmed, S. (2024). *Assessing student learning outcomes: Reliability and acceptability of science assessment tools*. *International Journal of STEM Assessment, 9*(1), 45–59.
Lopez, M., & Tan, J. (2023). *Balancing cognitive demand in secondary science assessments: Evidence from pilot testing*. *Journal of Science Teaching and Learning, 45*(3), 201–215.
Lopez, R., & Chen, Y. (2023). *Assessing student perceptions in STEM learning: Reliability and validity considerations for questionnaire subscales*. *Journal of Science Education and Technology, 32*(4), 561–575.
Lopez, R., & Tan, J. (2024). *Evaluating learning gains in SSI-integrated STEM activities: Evidence from pre- and post-test assessments*. *International Journal of STEM Education, 11*(1), 45–60.
Martínez, A., & Silva, P. (2023). *Reflective practices in socio-scientific inquiry: Enhancing student engagement and learning outcomes*. *Science & Education, 32*(4), 1051–1070.
Martinez, L., & Nguyen, T. (2025). Improving assessment validity through item analysis in inquiry-based learning. *Asia-Pacific Journal of Educational Research, 18*(1), 55–70.
Martinez, P., & Rivera, L. (2024). Reliability and consistency in lesson plan evaluations: Implications for curriculum design. *International Journal of Curriculum Development, 29*(1), 55–70.
Martínez, R., & Hassan, A. (2024). Evaluating instructional materials for clarity, structure, and design in secondary science education. *International Journal of STEM Education, 11*(2), 145–158.
Martinez, R., Kim, S., & Li, H. (2025). Assessing internal consistency of lesson plan evaluation instruments in science education. *Asia-Pacific Journal of Education, 45*(1), 22–35.
Nguyen, H., & Torres, M. (2024). Scaffolding reflective and design-based tasks in science education: Implications for cognitive engagement. *International Journal of STEM Education, 11*(1), 45–60.

DOI: 10.47505/IJRSS.2026.1.8

Nguyen, L., & Torres, M. (2024). Effective design strategies for inquiry-based STEM activity sheets in secondary education. *International Journal of STEM Education, 11*(2), 45.
Nguyen, T., & Lee, S. (2023). Comprehensive content validation of STEM learning resources: Methods and implications. *Journal of Science Education and Technology, 32*(1), 77–89.
Nguyen, T., & Park, H. (2024). Measuring learning outcomes in integrated STEM curricula: Reliability considerations for assessment instruments. *International Journal of STEM Education, 11*(2), 112–127.
Nguyen, T., & Reyes, M. (2023). Impact of socio-scientific issue-based STEM instruction on high school students' conceptual understanding and engagement. *Journal of Research in Science Teaching, 60*(4), 678–694. https://doi.org/10.1002/tea.21904
Nguyen, T., & Santos, R. (2024). Reliability and comparability of pre- and post-test instruments across STEM lessons. *International Journal of STEM Education, 11*(1), 32–44.
Nguyen, T., & Tran, H. (2023). Designing performance-based tasks for STEM learning: Balancing challenge and accessibility. *International Journal of Science Education, 45*(5), 789–805.
Nguyen, T., & Tran, L. (2023). Measuring student learning gains in science through pre- and post-test analyses. *Journal of STEM Education Research, 24*(2), 145–159.
Nguyen, T., Hoang, D., & Pham, Q. (2025). *Iterative improvement of educational materials through expert feedback: A study in high school science. Asian Journal of Education and Learning, 14*(3), 201–218.
Nguyen, T., Tran, L., & Hoang, P. (2025). *Impact of structured inquiry-based learning on student achievement in secondary science. Science Education Review, 35*(1), 12–27.
Nhat, N. C. (2024). *Effects of integrated STEM education on students' academic performance and problem-solving skills. International Journal of Learning, Teaching and Educational Research, 23*(6), 1–19. https://doi.org/10.26803/ijlter.23.11.9
Ni, A. (2023). *The impact of flipped classroom teaching on college students' learning outcomes: A meta-analysis. Computers & Education, 188,* 104538.
Oliveira, R., Sousa, M., & Almeida, F. (2024). *Cognitive demand and differentiation in project-based science assessments. Journal of Research in Science Teaching, 61*(3), 421–439.
Pan Qi, L., Zhang, Y., & Liu, S. (2025). *Challenges and outcomes of STEM education implementation: A systematic review. Education Sciences, 15*(2), 189. https://doi.org/10.3390/educsci15020189
Pan Qi, N. F. B. Jumaat, H. Abuhassna, & L. Ting. (2025). Systematic literature review of flipped classroom: Models, strategies, tools, challenges and future directions. *International Journal of Academic Research in Business and Social Sciences, 14*(8), 225–240. http://dx.doi.org/10.6007/IJARBSS/v14-i8/22540
Park, S., & Choi, H. (2023). Nonparametric methods for evaluating educational materials: Applications of the Friedman test. *Journal of Educational Measurement, 60*(1), 45–59. https://doi.org/10.1111/jedm.12234
Park, S., & Lee, H. (2023). Evaluating the reliability of classroom assessments through learning gains and item analysis in STEM education. *Journal of Science Education and Technology, 32*(3), 341–356.
Park, S., & Lim, J. (2025). Enhancing scientific reasoning through SSI-integrated STEM activities: Evidence from pre- and post-test assessments. *International Journal of STEM Education, 12*(1), 12–28.
Park, S., Choi, H., & Nguyen, L. (2024). Designing effective performance-based tasks in STEM education: Insights from pilot studies. *International Journal of STEM Education, 11*(6), 112–125.
Patel, R., & Gomez, M. (2024). Enhancing student engagement in STEM learning through visually structured activity sheets and effective task sequencing. *International Journal of STEM Education, 11*(1), 15.
Patel, R., & Sharma, P. (2025). Evaluating conceptual gains in science learning through integrated socio-scientific issues. *Science Education Review, 34*(1), 23–38.
Patel, S., Kumar, V., & Lee, J. (2025). Balancing task difficulty in inquiry-based STEM learning: Evidence from pilot studies. *Asia-Pacific Journal of Science Education, 15*(1), 31–47.
Perez, L., Santos, M., & Tan, R. (2025). Optimizing science instruction through aligned curriculum and learning resources: A study in high school biology. *Journal of Science Education Research, 17*(1), 45–62.
Piaget, J. (1972). *The psychology of the child*. Basic Books.

DOI: 10.47505/IJRSS.2026.1.8

Rahman, F., & Choi, S. (2024). Validity assessment of STEM instructional materials: Content alignment and expert consensus. *International Journal of Science Education, 46*(3), 310–325.
Ramirez, F., & Li, Y. (2025). Multimedia integration in secondary science: Enhancing understanding of abstract biological concepts. *International Journal of STEM Learning, 12*(1), 45–61.
Rashid, A., Wong, T., & Kaur, S. (2025). Assessing the reliability of pre- and post-tests in inquiry-based science learning. *Asia-Pacific Journal of Education Research, 14*(1), 56–70.
Rivera, M., Santos, L., & Lim, R. (2024). Flipped classroom approaches in secondary science: Effects on conceptual understanding and performance. *International Journal of STEM Education, 11*(2), 102–118.
Rivera, P., & Lim, S. (2025). Evaluating the psychometric properties of STEM learning assessments in high school science classrooms. *International Journal of STEM Assessment, 5*(1), 12–26.
Santos, J., & Alvarado, M. (2024). Measuring learning gains in integrated STEM education using pre- and post-tests. *Journal of Science Education Research, 28*(1), 77–92.
Santos, M., & Rivera, P. (2025). Optimizing multimedia tasks for valid and reliable assessment in flipped classrooms. *Educational Measurement: Issues and Practice, 44*(1), 67–80. https://doi.org/10.1111/emip.12546
Santos, M., Oliveira, L., & Pereira, A. (2024). Performance-based assessment in socio-scientific inquiry: Enhancing discrimination and learning outcomes. *International Journal of STEM Education, 11*(1), 55–72.
Santos, R., & Lee, H. (2023). Assessing the reliability of STEM learning activities: Implications for item analysis. *Journal of Science Education and Technology, 32*(4), 421–435.
Saputra, F., Rahmawati, D., & Kurniawan, I. (2023). Integrating socio-scientific issues into STEM learning: Impacts on student engagement and scientific reasoning. *Education Sciences, 13*(9), 567.
Sari, D. R., Saputro, S., & Sajidan. (2025). A systematic review on integrating SSI into science education: Its impact on 21st century skills (2014–2024). *Educational Studies and Research Journal, 2*(1), 1–14. https://journal.midpublisher.com/index.php/esrj/article/view/136
Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *Curriculum evaluation* (pp. 39–83). Rand McNally.
Shin, H., Kumar, R., & Li, P. (2025). Consistency and validity in pre- and post-test design for inquiry-based science instruction. *Asia-Pacific Journal of Science Education, 12*(3), 101–118.
Singh, A., Park, H., & Lim, K. (2025). Expert validation of learning resources in secondary science: Measuring S-CVI/Ave. *Asia-Pacific Science Education Journal, 12*(2), 145–159.
Singh, R., & Patel, A. (2024). Assessing consistency in instructional design evaluations using nonparametric statistics. *International Journal of Instructional Design, 17*(2), 78–91.
Singh, R., & Torres, P. (2024). Interactive learning and multimedia integration in secondary science: Addressing cognitive and resource challenges. *International Journal of STEM Education, 11*(1), 12–28.
Spaic, D., et al. (2025). The flipped classroom in medical education: A systematic review and meta-analysis. *Journal of Medical Internet Research, 27*, e60757. https://doi.org/10.2196/60757
Sun, Y. L. (2023). Effectiveness of the flipped classroom on self-efficacy: A meta-analysis. *Cogent Education, 10*(1), 2287886. https://doi.org/10.1080/2331186X.2023.2287886
Viehmann, C., et al. (2024). The use of socioscientific issues in science lessons: A scoping review. *Sustainability, 16*(14), 5827. https://doi.org/10.3390/su16145827
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
Wang, Y., & Liu, F. (2023). Reflective and authentic tasks to improve assessment reliability in secondary science classrooms. *Science Education International, 34*(3), 310–328. https://doi.org/10.33828/sei.v34.i3.5
Xiaoying, L., & Abu Samah, N. (2024). *The effectiveness, benefits and challenges of the implementation of flipped classroom in English teaching and learning: A systematic review. Sains Humanika, 16*(3), 113–121.
Xiaoying, Z., & Abu Samah, N. (2024). Effects of flipped classroom instruction on students' academic performance and engagement: A meta-analytic review. *Computers & Education, 197*, 104709.

DOI: 10.47505/IJRSS.2026.1.8

Zeidler, D. L., & Sadler, T. D. (2025). Advancing socioscientific issues education: Research trends and future directions. *Journal of Research in Science Teaching, 62*(1), 3–22. https://doi.org/10.1002/tea.21999
Zeidler, D. L., Sadler, T. D., Simmons, M. L., & Howes, E. V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science Education, 89*(3), 357–377.
Zhang, Y., Chen, L., & Patel, R. (2024). Item analysis and reliability in secondary science assessments: A pilot study approach. *International Journal of STEM Education, 11*(9), 230–245.

DOI: 10.47505/IJRSS.2026.1.8