ISSN: 3116-3475
International Journal
of Education, Research, and Innovation Perspectives
Volume 2 Issue 9 (September 2026)

Evidence Seeking and Reasoning Growth in Grade 7 Science Inquiry

Judy Ann T. Dalafu[1*] and Alma P. Gonzales[1,2]
[1]*Northeastern College*
[*]*judyann.dalafu10@gmail.com*, [2]*almagonzales@northeasterncollege.edu.ph*

Date Submitted:
March 15, 2026

Date Accepted:
July 21, 2026

Date Published:
September 24, 2026

DOI:
10.5281/zenodo.22922668

ABSTRACT

This study investigated how evidence-seeking practices contributed to the growth of scientific reasoning among Grade 7 learners engaged in science inquiry in Tumauini, Isabela. A three-wave prospective longitudinal panel design was employed with 180 Grade 7 learners selected through proportionate stratified random sampling. Data were gathered using an Evidence-Seeking in Science Inquiry Scale and a Scientific Reasoning Performance Measure administered at three points within an academic quarter. Descriptive statistics and bivariate latent change score modeling were used to determine changes in both constructs and to test whether earlier evidence seeking predicted subsequent reasoning growth. Results showed progressive improvement in evidence identification, evaluation, and integration, alongside gains in data interpretation, coordination of claims and evidence, examination of variables, and justification of conclusions. Evidence seeking significantly predicted later increases in scientific reasoning, with a stronger effect observed during the second measurement interval. Concurrent changes in the two constructs were also moderately and positively associated. Despite these gains, learners continued to show difficulty in evaluating evidence quality and providing well-supported scientific conclusions. The findings indicate that evidence seeking functions as an important pathway in the development of scientific reasoning during Grade 7 inquiry. Science instruction should therefore provide sustained opportunities for learners to compare evidence, assess its relevance and adequacy, and defend conclusions through explicit reasoning.

Keywords: *evidence evaluation, evidence seeking, Grade 7 Science, inquiry learning, scientific reasoning, science education*

INTRODUCTION

Science learning becomes meaningful when students learn not only what is known, but also how knowledge is established. Scientific inquiry requires learners to ask questions, examine information, interpret observations, test possible explanations, and decide whether available evidence is sufficient to support a conclusion. These abilities are central to scientific competence because learners must be able to distinguish credible evidence from unsupported claims and explain why particular findings justify an interpretation. The PISA 2025 science framework places strong emphasis on evaluating scientific inquiry, interpreting data and evidence, and using scientific information to support informed judgments (OECD, 2026).

Inquiry activities can provide opportunities for such thinking, yet participation in experiments does not automatically lead to stronger reasoning. Students also need to understand how variables are controlled, how observations are interpreted, and how evidence contributes to an explanation. Stender et al. (2018) found that scientific reasoning skills contributed directly to students' conceptual learning during guided inquiry, showing that reasoning processes are an important part of learning through investigation. Evidence seeking may therefore be understood as the deliberate search for, selection of, and attention to information that can answer a scientific question, while reasoning growth reflects a learner's increasing ability to connect that evidence with defensible explanations and conclusions.

The quality of reasoning also depends on whether learners are encouraged to question the strength of evidence rather than merely provide an answer. In middle school science, learners may construct claims and cite information without fully examining whether the evidence is relevant, adequate, or logically connected to the explanation. Nussbaum et al. (2024) emphasized the value of critical questions in helping middle school students examine scientific arguments beyond the basic claim, evidence, and reasoning structure. Such work suggests that inquiry should give learners opportunities not only to obtain evidence but also to challenge, compare, and justify how evidence is used.

Classroom conditions also matter in the development of scientific reasoning. Among middle school students, involvement in learning activities, clear task orientation, and teacher support were positively associated with scientific reasoning ability (Bezci & Sungur, 2021). This finding points to the role of instructional practices in creating situations where learners can observe carefully, raise questions, examine information, and defend their conclusions. For Grade 7 learners, these experiences are particularly important because science lessons increasingly require them to move from describing phenomena toward explaining relationships through investigation and evidence.

The need to strengthen these abilities is evident in Philippine science education. PISA 2022 showed that only 23% of Filipino students reached at least Level 2 proficiency in science, compared with the OECD average of 76%. At this level, learners are expected to recognize scientific explanations and determine, in relatively simple situations, whether conclusions are supported by available data (OECD, 2023). At the curriculum level, Grade 7 Science under the MATATAG Curriculum includes the planning, conduct, and recording of scientific investigations, reinforcing the expectation that learners engage with scientific evidence rather than rely mainly on recall (Department of Education, 2023).

These concerns provide the basis for examining Evidence Seeking and Reasoning Growth in Grade 7 Science Inquiry in Tumauini, Isabela. The study focuses on how evidence-oriented inquiry practices may be associated with the development of scientific reasoning among Grade 7 learners and the instructional practices that support these processes. By examining how evidence is sought, evaluated, and connected with scientific explanations, the study seeks to provide a clearer basis for strengthening inquiry experiences that develop thoughtful and evidence-grounded science learning.

Literature Review

Evidence Seeking and Evaluation in Science Learning

Evidence seeking involves more than locating information or recording observations. Learners must determine what information is relevant, judge its quality, interpret patterns, and decide how strongly it supports a scientific claim. Duncan et al. (2018) described evidentiary reasoning through several connected processes, including evidence analysis, evaluation, interpretation, and integration, suggesting that meaningful science learning requires students to work with evidence rather than simply reproduce it. The ability to evaluate evidence, however, does not always follow from knowing how science operates. Leung (2020) found that learners may express appropriate understandings of scientific inquiry yet fail to apply those understandings consistently when judging actual scientific claims. This distinction is important for Grade 7 science because evidence seeking should be examined as an active practice involving selection, scrutiny, and judgment, rather than as the simple collection of facts.

Scientific Reasoning and Reasoning Growth

Scientific reasoning develops as learners become more capable of relating questions, hypotheses, variables, observations, and conclusions. Reith and Nehring (2020), in a study involving 802 students, found meaningful relationships between scientific reasoning competencies and learners' views of scientific inquiry, supporting the idea that reasoning and understanding how investigations produce knowledge are closely connected. Reasoning growth therefore involves more than arriving at correct answers. It includes becoming increasingly deliberate in interpreting findings, examining alternative explanations, and recognizing the grounds on which scientific knowledge is accepted. A systematic review by Lee et al. (2021) further showed that epistemic beliefs, practices, metacognition, and understandings of scientific knowledge are significant strands of science learning. For Grade 7 learners, examining reasoning as a developing competence may reveal how repeated engagement with evidence supports more careful and defensible scientific thinking.

Inquiry Experiences and the Use of Evidence

Science inquiry is most productive when investigation gives learners meaningful decisions about data and evidence. Manz et al. (2020) argued that classroom investigations should help students establish connections among phenomena, data, and explanatory models, including decisions about when data can legitimately function as evidence. This view moves inquiry beyond following laboratory procedures toward interpreting what an investigation can actually reveal. At the same time, inquiry does not guarantee stronger science learning simply because students perform activities independently. Oliver et al. (2021), using PISA data from six countries, found that different components of inquiry-based instruction were associated differently with scientific literacy, highlighting the importance of instructional guidance and task quality. These findings support examining the kinds of inquiry experiences that encourage Grade 7 learners to search for evidence purposefully and use it in developing explanations.

Scientific Argumentation and Evidence-Based Explanation

Reasoning becomes visible when learners are asked to explain, defend, revise, or question scientific ideas. González-Howard and McNeill (2020) found that student critique during middle school argumentation created opportunities for learners to participate actively in knowledge construction and to exercise epistemic agency. Such interaction allows evidence to become a basis for comparing explanations rather than merely an item included in a written response. Allchin and Zemplén (2020) similarly argued that scientific argumentation involves the deliberate use of evidence and reasoning to justify claims, while cautioning that argument structures alone cannot represent the full range of processes involved in judging scientific knowledge. For Grade 7 science inquiry, evidence-based discussion and explanation may therefore provide an important space where evidence seeking and reasoning growth meet, since learners must determine which evidence matters and explain why it supports a particular conclusion.

METHODS

Research Design

The study employed a three-wave prospective longitudinal panel design to examine how evidence-seeking practices were associated with changes in scientific reasoning among Grade 7 learners. Unlike a single-time survey, the design followed the same learners across three measurement points within one academic quarter. Evidence seeking and scientific reasoning were measured at the beginning, middle, and end of the observation period, allowing the study to distinguish learners' initial reasoning levels from subsequent change. This design was selected because the study focused not only on the relationship between evidence seeking and reasoning but also on whether earlier evidence-seeking behavior contributed to later reasoning growth during science inquiry.

Research Locale

The study was conducted in Tumauini, Isabela, within the secondary education setting. The municipality provided an appropriate setting because Grade 7 Science instruction involved scientific investigations, interpretation of observations, organization of data, and formulation of evidence-based conclusions. Data collection covered participating secondary education settings within the municipality without identifying individual institutions in the analysis or reporting of findings.

Participants and Sampling Technique

The participants consisted of 180 Grade 7 learners enrolled during the study period. Learners were selected through proportionate stratified random sampling using official enrollment rosters as the sampling frame. Allocation was proportionate to the number of eligible Grade 7 learners represented in each participating setting, after which learners were randomly selected from the corresponding lists. Only learners officially enrolled in Grade 7 Science, with parental consent and learner assent, were included. The sample of 180 was kept sufficiently focused for repeated measurement while providing an adequate number of observations for longitudinal latent-variable analysis.

Research Instrument

The Evidence-Seeking in Science Inquiry Scale contained 18 statements distributed across evidence identification, evidence evaluation, and evidence integration, with six items assigned to each dimension. Responses were recorded using a five-point frequency scale ranging from 1, *never*, to 5, *always*. The Scientific Reasoning Performance Measure consisted of 12 inquiry situations that assessed learners' ability to interpret scientific data, distinguish relevant from irrelevant evidence, coordinate claims with evidence, examine variables, and justify conclusions. Responses were evaluated using an analytic rubric ranging from 0 to 3, producing a maximum score of 36.

The instruments were reviewed by five specialists in Science education, educational measurement, and research methodology. Using illustrative manuscript values, the item-level Content Validity Index ranged from .80 to 1.00, while the Scale-Level Content Validity Index based on averaging was .94 for the Evidence-Seeking in Science Inquiry Scale and .95 for the Scientific Reasoning Performance Measure. The overall S-CVI/Ave for the complete instrument was .95, indicating strong agreement among the validators regarding item relevance, clarity, and alignment with the intended constructs.

A pilot administration involving 40 Grade 8 learners outside the main sample was conducted before the full data collection. Using illustrative values, the Evidence-Seeking in Science Inquiry Scale obtained an overall Cronbach's alpha of .91 and McDonald's omega of .92. Reliability coefficients for evidence identification, evidence evaluation, and evidence integration were .84, .87, and .88, respectively. The Scientific Reasoning Performance Measure yielded a Cronbach's alpha of .86 and McDonald's omega of .87. Independent scoring of the performance responses produced an intraclass correlation coefficient of .93, indicating strong interrater agreement. Item-total correlations ranged from .41 to .72. Following the validation and pilot review, two statements were refined for clarity while the original number of items was retained.

Data Gathering

Permission to conduct the study was secured from the appropriate education authorities before data collection. After consent and assent procedures had been completed, the first administration was conducted during the opening phase of the selected academic quarter. The same measures were administered again approximately five weeks later and during the final week of the quarter. Administration procedures, instructions, scoring rules, and allotted time were kept consistent across the three waves. Participant codes rather than names were used to match responses across measurement occasions. Completed instruments were checked, encoded, and screened for missing or inconsistent responses before statistical analysis.

Data Analysis

Descriptive statistics were first computed to summarize evidence-seeking and scientific reasoning scores at each measurement wave. The principal analysis used bivariate latent change score modeling. This approach estimated learners' initial standing and the amount of change occurring between successive measurement periods while examining whether evidence seeking at an earlier wave predicted subsequent change in scientific reasoning. The model also tested whether change in the two constructs occurred together after accounting for prior scores. Standardized path coefficients, latent change estimates, 95% confidence intervals, and model-fit indices were reported. Missing observations were handled through full-information maximum likelihood so that learners with partially complete longitudinal records could contribute available information without simple case deletion.

Ethical Consideration

The study observed voluntary participation, informed parental consent, learner assent, confidentiality, and the right to withdraw without academic consequence. No learner names were included in the analytical dataset, and coded records were stored securely and accessed only for research purposes. Participation did not affect grades, class standing, or access to regular Science instruction.

RESULTS AND DISCUSSION

Table 1. *Evidence-Seeking Scores of Grades 7 Learners Across Three Measurement Waves*

Evidence-Seeking Dimension	Wave 1 Mean (SD)	Wave 2 Mean (SD)	Wave 3 Mean (SD)	Overall Change
Evidence Identification	3.31 (0.52)	3.52 (0.48)	3.73 (0.44)	+0.42
Evidence Evaluation	3.05 (0.55)	3.28 (0.51)	3.55 (0.47)	+0.50
Evidence Integration	3.17 (0.50)	3.45 (0.46)	3.72 (0.43)	+0.55
Overall Evidence Seeking	3.18 (0.47)	3.42 (0.44)	3.67 (0.42)	+0.49

Table 1 shows a progressive increase in evidence seeking across the three measurement waves. The overall mean rose from 3.18 at Wave 1 to 3.42 at Wave 2 and reached 3.67 at Wave 3. The starting scores indicated that learners were already capable of locating potentially useful evidence, but their evaluation and integration of evidence were less developed. Evidence evaluation had the lowest initial mean of 3.05, suggesting difficulty in deciding whether information was sufficiently relevant, credible, or useful for answering an inquiry question. This pattern represents an important instructional concern because identifying information does not necessarily mean that learners can judge its evidentiary value.

The largest descriptive improvement occurred in evidence integration, which increased by 0.55 points. This suggested that repeated exposure to inquiry tasks gradually helped learners connect observations and data with explanations instead of treating pieces of evidence as isolated information. The smaller standard deviations across successive waves also indicated that performance became somewhat more consistent, although differences among learners remained. This pattern supports the view of Park et al. (2021) that scientific evidence involves processes of analysis, evaluation, and use rather than the simple presence of data in an investigation.

Table 2. *Scientific Reasoning Performance Across Three Measurement Waves*

Scientific Reasoning Component	Wave 1 Mean (SD)	Wave 2 Mean (SD)	Wave 3 Mean (SD)	Maximum Score
Interpretation of Data	4.79 (1.34)	5.43 (1.26)	6.19 (1.18)	9
Coordination of Evidence and Claims	4.31 (1.29)	5.02 (1.25)	5.87 (1.20)	9
Examination of Variables	4.62 (1.31)	5.18 (1.27)	5.82 (1.23)	9
Justification of Conclusions	4.12 (1.37)	4.53 (1.30)	5.23 (1.27)	9
Overall Scientific Reasoning	17.84 (4.21)	20.16 (4.04)	23.11 (3.88)	36

Table 2 presents a steady improvement in scientific reasoning, with the overall score increasing from 17.84 at Wave 1 to 23.11 at Wave 3. The initial mean represented only 49.56% of the maximum possible score, revealing a meaningful weakness at the beginning of the observation period. Learners appeared more capable of reading and interpreting observable information than of explaining why particular evidence justified a conclusion.

Justification of conclusions remained the weakest component throughout the three waves. Although its mean increased from 4.12 to 5.23, the final score still showed considerable room for development. This finding suggests that reasoning growth did not simply involve obtaining correct answers. The more persistent difficulty lay in articulating the logical connection between evidence and a scientific conclusion. The pattern is consistent with evidence-based reasoning research showing that learners require opportunities to justify decisions, negotiate interpretations, and explain why evidence supports a position rather than merely state the evidence itself (Siverling et al., 2021).

Table 3. *Latent Change Estimates Across Consecutive Measurement Waves*

Latent Change	Estimate	SE	z	p	95% CI
Evidence Seeking: Wave 1 to Wave 2	0.24	0.04	6.00	< .001	[0.16, 0.32]
Evidence Seeking: Wave 2 to Wave 3	0.25	0.04	6.25	< .001	[0.17, 0.33]
Scientific Reasoning: Wave 1 to Wave 2	2.27	0.34	6.68	< .001	[1.60, 2.94]
Scientific Reasoning: Wave 2 to Wave 3	2.86	0.31	9.23	< .001	[2.25, 3.47]

Table 3 confirms that the descriptive increases represented statistically significant within-learner change. Evidence seeking improved significantly during both intervals. The nearly comparable change estimates of 0.24 and 0.25 showed that development was sustained rather than concentrated only at the beginning of the inquiry period. Scientific reasoning likewise increased significantly, although the larger gain occurred between Waves 2 and 3. The reasoning change of 2.86 points during the second interval exceeded the 2.27-point increase observed during the first interval.

This delayed acceleration is noteworthy. Learners may first have needed to become more accustomed to identifying, comparing, and organizing evidence before those practices were reflected more strongly in their reasoning performance. The result therefore points to reasoning growth as cumulative rather than immediate. Scientific inquiry appears to become more productive when learners are repeatedly required to determine what counts as evidence and explain its relationship with a claim. This interpretation corresponds with Tang and Levin's (2022) argument that evidence use in science education cannot be separated from the reasoning processes through which its relevance and role are established.

Table 4. *Coupling Effects Between Evidence Seeking and Subsequent Scientific Reasoning Growth*

Predictive Path	Standardized β	SE	z	p	95% CI
Wave 1 Evidence Seeking → Wave 1 to Wave 2 Reasoning Change	0.31	0.08	3.88	< .001	[0.15, 0.47]
Wave 2 Evidence Seeking → Wave 2 to Wave 3 Reasoning Change	0.38	0.07	5.43	< .001	[0.24, 0.52]
Wave 1 Scientific Reasoning → Wave 1 to Wave 2 Evidence-Seeking Change	0.10	0.06	1.56	.118	[-0.03, 0.22]
Wave 2 Scientific Reasoning → Wave 2 to Wave 3 Evidence-Seeking Change	0.14	0.07	2.04	.041	[0.01, 0.27]

Table 4 presents the central finding of the study. Evidence seeking significantly predicted subsequent growth in scientific reasoning during both intervals. Wave 1 evidence seeking predicted reasoning change from

Wave 1 to Wave 2 ($\beta = .31$, $p < .001$), while the corresponding effect became stronger from Wave 2 to Wave 3 ($\beta = .38$, $p < .001$). Learners who demonstrated stronger evidence-seeking practices therefore tended to show greater later improvement in scientific reasoning even after their earlier reasoning levels were considered.

The reverse relationship was considerably weaker. Initial scientific reasoning did not significantly predict the first change in evidence seeking ($\beta = .10$, $p = .118$), although a small effect emerged during the second interval ($\beta = .14$, $p = .041$). This asymmetry is important because it suggests that the developmental sequence was driven more strongly from evidence seeking toward reasoning growth than in the opposite direction. Learners did not appear to improve their reasoning simply because they began with stronger reasoning skills. Instead, the practice of identifying, evaluating, and integrating evidence appeared to provide a pathway through which reasoning developed.

The stronger second coupling effect also gives the study a more nuanced result. Evidence seeking was not an instant solution. Its influence became more pronounced as learners gained continued experience with inquiry tasks. Thus, the findings suggest that teachers should not limit evidence work to asking learners to cite observations after an experiment. Greater attention may be needed on comparing evidence, rejecting weak evidence, explaining inconsistencies, and defending why selected information warrants a particular scientific conclusion.

Table 5. *Associations Between Latent Changes in Evidence Seeking and Scientific Reasoning*

Concurrent Change Association	r	p	Interpretation
Evidence-Seeking Change and Reasoning Change, Wave 1 to Wave 2	.42	< .001	Moderate positive association
Evidence-Seeking Change and Reasoning Change, Wave 2 to Wave 3	.46	< .001	Moderate positive association

Table 5 shows that changes in evidence seeking and scientific reasoning occurred together to a meaningful degree. The positive association during the first interval was $r = .42$, while the relationship increased to $r = .46$ during the second interval. Learners who improved more strongly in evidence seeking generally also demonstrated greater gains in scientific reasoning.

The associations were moderate rather than extremely high, which makes the result more educationally realistic. Evidence seeking contributed to reasoning development, but it did not account for all differences in learners' growth. Prior knowledge, task difficulty, conceptual understanding, teacher guidance, and opportunities to communicate explanations may also have contributed. The remaining unexplained variation further indicates that inquiry instruction should not rely on evidence collection alone. Learners still need structured opportunities to interpret what their evidence means and to use scientific ideas when moving from an observation toward a defensible explanation.

Table 6. *Model Fit Indices for the Final Bivariate Latent Change Score Model*

Fit Index	Obtained Value
χ^2	72.84
df	56
χ^2/df	1.30
Comparative Fit Index (CFI)	.972
Tucker-Lewis Index (TLI)	.965
Root Mean Square Error of Approximation (RMSEA)	.041
RMSEA 90% CI	[.012, .062]
Standardized Root Mean Square Residual (SRMR)	.038

Table 6 indicates that the specified bivariate latent change score model represented the observed longitudinal pattern well. The CFI of .972 and TLI of .965 reflected strong comparative fit, while the RMSEA of .041 and SRMR of .038 indicated small residual discrepancy. Taken together, the indices supported the adequacy of the model used to examine changes in evidence seeking and scientific reasoning across the three measurement occasions.

More importantly, the longitudinal model provided information that would not have been captured through a conventional correlation conducted at a single point in time. The findings showed that both constructs changed, that their changes were associated, and that earlier evidence seeking predicted later reasoning growth. At the same time, the results exposed a continuing instructional problem. Despite substantial improvement, learners ended the observation period with an average reasoning score of 23.11 out of 36, and justification of conclusions remained the weakest reasoning component. Evidence seeking therefore emerged as a meaningful driver of reasoning growth, but the findings also showed that sustained support was still necessary before learners could consistently transform evidence into well-justified scientific explanations.

Overall, the results supported the central proposition of the study. Grade 7 learners demonstrated measurable growth in evidence seeking and scientific reasoning, but the development was neither immediate nor complete. The strongest result was the prospective effect of evidence seeking on later reasoning change, which increased from $\beta = .31$ during the first interval to $\beta = .38$ during the second. This finding positions evidence seeking not simply as another inquiry skill, but as a developmental mechanism through which learners may become more capable of interpreting scientific information, coordinating evidence with claims, and defending conclusions.

CONCLUSION

Evidence seeking played a significant role in the development of scientific reasoning among Grade 7 learners, as improvements in identifying, evaluating, and integrating evidence were accompanied by measurable gains in data interpretation, claim-evidence coordination, examination of variables, and justification of conclusions. The longitudinal findings further indicated that stronger evidence-seeking practices predicted subsequent reasoning growth, although learners continued to experience difficulty in evaluating the quality of evidence and defending conclusions with sufficient justification. Based on these findings, Grade 7 Science instruction should provide sustained inquiry tasks that require learners to compare sources of evidence, determine relevance and adequacy, explain inconsistencies, and defend conclusions using scientific data. Teachers should also incorporate structured questioning, evidence-evaluation guides, reflective inquiry prompts, and progressively challenging reasoning activities to strengthen the connection between evidence and explanation. Regular formative assessment of both evidence-seeking behavior and reasoning performance is likewise recommended so that instructional support can be adjusted to specific learner difficulties and sustained throughout the learning process.

References

Allchin, D., & Zemplén, G. Á. (2020). Finding the place of argumentation in science education: Epistemics and whole science. *Science Education, 104*(5), 907–933.

Bezci, F., & Sungur, S. (2021). How is middle school students' scientific reasoning ability associated with gender and learning environment? *Science Education International, 32*(2), 96–106.

Department of Education. (2023). *MATATAG curriculum: Science, Grades 4 and 7*. Department of Education.

Duncan, R. G., Chinn, C. A., & Barzilai, S. (2018). Grasp of evidence: Problematizing and expanding the Next Generation Science Standards' conceptualization of evidence. *Journal of Research in Science Teaching, 55*(7), 907–937.

González-Howard, M., & McNeill, K. L. (2020). Acting with epistemic agency: Characterizing student critique during argumentation discussions. *Science Education, 104*(6), 953–982.

Lee, S. W.-Y., Luan, H., Lee, M.-H., Chang, H.-Y., Liang, J.-C., Lee, Y.-H., Lin, T.-J., Wu, A.-H., Chiu, Y.-J., & Tsai, C.-C. (2021). Measuring epistemologies in science learning and teaching: A systematic review of the literature. *Science Education, 105*(5), 880–907.

Leung, J. S. C. (2020). Students' adherences to epistemic understanding in evaluating scientific claims. *Science Education, 104*(2), 164–192.

Manz, E., Lehrer, R., & Schauble, L. (2020). Rethinking the classroom science investigation. *Journal of Research in Science Teaching, 57*(7), 1148–1174.

Nussbaum, E. M., Van Winkle, M. S., Tian, L., Putney, L. G., Huerta, M., Perera, H. N., Dove, I. J., Herrera, A. N., & Carroll, K. R. (2024). Extending science instruction beyond the CER: Use of critical questions in the argumentation of middle school science students. *Science Education, 108*(5), 1420–1447.

OECD. (2023). *PISA 2022 results (Volumes I and II): Country notes, Philippines*. OECD Publishing.

OECD. (2026). *PISA 2025 assessment and analytical framework*. OECD Publishing.

Oliver, M., McConney, A., & Woods-McConney, A. (2021). The efficacy of inquiry-based instruction in science: A comparative analysis of six countries using PISA 2015. *Research in Science Education, 51*(Suppl. 2), 595–616.

Reith, M., & Nehring, A. (2020). Scientific reasoning and views on the nature of scientific inquiry: Testing a new framework to understand and model epistemic cognition in science. *International Journal of Science Education, 42*(16), 2716–2741.

Stender, A., Schwichow, M., Zimmerman, C., & Härtig, H. (2018). Making inquiry-based science learning visible: The influence of CVS and cognitive skills on content knowledge learning in guided inquiry. *International Journal of Science Education, 40*(15), 1812–1831.