JPPIPA 10(9) (2024)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Basic Science Concept Mastery among Prospective Elementary Teachers: A Descriptive Baseline for Diagnostic Assessment

Fonny Katili[1]*, Citra Raflesia[1], Yusak Ratunguri[1]

[1] Department of Elementary School Teacher Education, Universitas Negeri Manado, Tondano, Indonesia

Received: July12, 2024
Revised: August21, 2024
Accepted: September25, 2024
Published: September30, 2024

Corresponding Author:
Fonny Kotili
fonny_katili@unima.ac.id

DOI: 10.29303/jppipa.v10i9.16023

Abstract: Aggregate achievement scores are frequently used to summarize prospective elementary teachers' science knowledge, yet such scores can obscure the conceptual patterns that matter for teacher preparation. This study examined the distribution and interpretive limits of basic science achievement among 80 prospective elementary teachers enrolled in a Basic Science Concepts course. A descriptive quantitative cross-sectional design was used, and archived total test scores on a 0-100 scale were analyzed using measures of central tendency, dispersion, category frequencies, and a 95% confidence interval. The mean score was 77.88 (SD = 3.23; 95% CI = 77.16-78.59), with 81.25% of students classified as good, 3.75% as very good, and 15.00% as adequate. The narrow range and small dispersion indicate a relatively homogeneous level of overall achievement. However, because the archived dataset contained only total scores and no item-level responses, reasoning, or confidence data, the study cannot identify which science domains were most difficult or diagnose misconceptions. The contribution of the study is therefore a cautious baseline: aggregate scores can support cohort-level monitoring but are insufficient for diagnostic claims. Teacher-education assessment should complement total scores with item-level, multi-tier, and reasoning-based measures to guide targeted conceptual reinforcement.

Keywords: Basic science; Concept mastery; Diagnostic assessment; Elementary teacher education; Prospective teacher; Science education

Introduction

Elementary teachers are expected to introduce learners not only to scientific facts but also to explanations, evidence, inquiry, and the relationships among scientific ideas. This expectation makes science content knowledge a central component of elementary teacher preparation because inaccurate or fragmented content knowledge can shape the examples teachers select, the explanations they accept, and the questions they ask in class. Research on prospective elementary teachers has therefore treated science understanding as an important foundation for later pedagogical decision making, alongside self-efficacy and knowledge of how students think about science (Bleicher & Lindgren, 2005; Menon & Sadler, 2016; Sabel, Forbes, & Zangori, 2015).

A persistent challenge is that prospective elementary teachers enter teacher education with heterogeneous science experiences and may retain alternative conceptions from prior schooling or everyday reasoning. Domain-specific studies have documented difficulties in astronomy, physical science, and life science, including the causes of day and night and seasons, physical-science phenomena, and the relationship between photosynthesis and cellular respiration (Atwood & Atwood, 1995, 1996; Brown & Schwartz, 2009; Burgoon, Heddle, & Duran, 2011). These findings are consequential because a teacher can obtain

How to Cite:
Katili, F., Raflesia, C., & Ratunguri, Y. (2024). Basic Science Concept Mastery among Prospective Elementary Teachers: A Descriptive Baseline for Diagnostic Assessment. *Jurnal Penelitian Pendidikan IPA*, *10*(9), 7209–7216. https://doi.org/10.29303/jppipa.v10i9.16023

a generally acceptable course score while still holding robust alternative explanations in particular domains.

Science teacher-education research also shows that well-designed coursework can improve prospective teachers' science learning, confidence, and willingness to teach science. Constructivist and inquiry-oriented experiences, hands-on investigation, reflection, and formative assessment have been associated with gains in conceptual understanding, science-teaching self-efficacy, attitudes, and the capacity to use students' thinking instructionally (Avery & Meyer, 2012; Bleicher & Lindgren, 2005; Menon & Azam, 2021; Riegle-Crumb et al., 2015; Sabel et al., 2015). A systematic review likewise indicates that inquiry-based science education in teacher preparation can support both learning of science and learning how to teach through inquiry (Strat, Henriksen, & Jegstad, 2024).

At the same time, the measurement used to represent 'concept mastery' matters. Conventional total scores efficiently summarize performance, but they collapse performance across items and domains into a single number. Conceptual-change research emphasizes that scientific understanding is not equivalent to the accumulation of correct answers; learners may answer correctly through memorized rules, partial reasoning, or guessing, while persistent alternative conceptions are best identified through evidence about both responses and reasoning (Duit & Treagust, 2003; Posner, Strike, Hewson, & Gertzog, 1982). Diagnostic-assessment research therefore recommends two-tier, three-tier, or related instruments that combine an answer with a reason and, in some designs, confidence information (Gurel, Eryilmaz, & McDermott, 2015; Treagust, 1988).

This distinction is especially relevant for prospective elementary teachers. A broad Basic Science Concepts course commonly spans matter, force and motion, energy, living systems, ecosystems, and Earth-space science. An overall course-test score can indicate whether a cohort is performing at an expected level, but it cannot show whether strength in one domain compensates for weakness in another, whether an incorrect response reflects a misconception or incomplete knowledge, or whether students can explain a concept in a form suitable for elementary learners. Recent work on multi-tier diagnostic assessment illustrates how reasoning information improves the defensibility of claims about conceptual difficulty and misconception (Nasir, Cari, Sunarno, & Rahmawati, 2024; Taşçı, 2024).

Previous studies provide a strong basis for expecting that prospective teachers' science knowledge can improve through teacher-education experiences, but the literature also reveals two recurring limitations. First, many studies focus on self-efficacy, attitudes, inquiry readiness, teaching practice, or a single science domain rather than reporting a transparent profile of broad basic-science achievement. Second, studies that diagnose misconceptions typically require item-level evidence and specialized instruments, whereas routine program data are often available only as total scores. Consequently, there remains a practical and methodological need to distinguish what aggregate achievement data can legitimately show from what requires a dedicated diagnostic assessment (Kruse et al., 2021; Menon, 2020; Norris, Morris, & Lummis, 2018; Seung, Park, Kite, & Choi, 2023).

In Indonesia, this distinction has direct implications for elementary teacher education because science courses must prepare future teachers to connect scientific ideas with everyday phenomena while avoiding oversimplified explanations. A cohort-level achievement profile can be useful for program monitoring and for identifying whether additional support is warranted, but the interpretation must remain proportional to the evidence available. Treating a total score as if it were a map of students' reasoning risks overstating the validity of the conclusions and may lead to poorly targeted remediation.

Accordingly, the scientific problem addressed in this study is not the identification of specific misconceptions from insufficient data, but the interpretation of aggregate basic-science achievement as a baseline for subsequent diagnostic assessment. The study asks two questions: (1) What is the level and distribution of basic science achievement among prospective elementary teachers in the observed cohort? and (2) What can, and what cannot, be inferred from the aggregate score pattern for planning conceptual reinforcement in teacher education? By answering these questions, the study offers a bounded descriptive contribution and explicitly separates cohort-level achievement evidence from item-level diagnostic claims.

The study was conducted in the Elementary School Teacher Education Program at Universitas Negeri Manado as the empirical setting for this broader assessment problem. The institutional context is therefore treated as the source of the data rather than as the sole contribution of the study. The interpretation is intended to inform similar teacher-education settings in which archived assessment data consist primarily of total scores and where decisions must be made about whether more diagnostic assessment is needed.

The purpose of the study is to describe the overall distribution of basic science achievement and to establish the interpretive boundary of the available data. The analysis does not claim to determine which science concepts are most difficult or to estimate a misconception rate because those questions require item-level responses, reasons, and preferably confidence data.

Method

Research Design

This study used a descriptive quantitative cross-sectional design based on archived course-assessment data. The design was selected to characterize the distribution of achievement in one cohort at a defined point in teacher preparation. No instructional treatment was manipulated and no causal comparison was intended.

Participants and Setting

The study was conducted in the Elementary School Teacher Education Program at Universitas Negeri Manado. Participants were 80 prospective elementary teachers who had completed or were taking the Basic Science Concepts course and for whom complete total-score data were available. Participation in the research use of the data was voluntary. Because the sample came from one program and was not randomly drawn from a wider population, generalization beyond the observed cohort should be made cautiously.

Instrument and Measurement Boundary

Concept mastery was represented by the total score from the Basic Science Concepts achievement assessment used in the course, reported on a 0-100 scale. The course assessment framework covered six content areas relevant to elementary science education: matter and its changes, force and motion, energy and its transformations, living systems, ecosystems, and Earth-space science. For the dataset available to this study, only the final total score for each participant was archived. Item-level responses, reasoning data, confidence ratings, the number and format of individual items, item difficulty and discrimination indices, reliability coefficients, and the original expert-validation records were not available for reanalysis. Accordingly, the assessment is treated here as a course achievement measure rather than as a psychometrically established diagnostic test.

This reporting decision is important for construct validity. The study therefore avoids interpreting total scores as direct evidence of domain-specific mastery or misconception. The six content areas in Table 1 describe the intended curricular scope of the assessment, but the available dataset does not permit separate scores for those domains. Future publication-level diagnostic claims would require recovery of the original instrument and item-level data, followed by content-validation and psychometric analysis.

Table 1. Scope of Basic Science Concepts Assessed

Content Area	Conceptual Scope
Matter and changes	properties of matter; physical and chemical change
Force and motion	force, motion, friction, and simple machines
Energy	forms, transfer, and transformation of energy
Living systems	structure, function, and life processes
Ecosystems	biotic-abiotic relationships and food webs
Earth and space	Earth processes, rotation, revolution, and celestial phenomena

Data Collection and Ethics

Data were obtained during scheduled course activities after permission from the study program and informed consent from participants. Student identities were replaced with numerical codes before analysis. Participation was voluntary, confidentiality was maintained, and participation did not affect course grades. Only de-identified total-score data were used in the analysis reported here.

Operational Definition

For this study, overall basic science achievement was operationalized as the total percentage score on the course assessment. The cohort mean represents the average proportion of the maximum score achieved. The complement of the mean (100 - 77.88 = 22.12 percentage points) is reported only as an unearned-score proportion, meaning the average distance from the maximum possible score. It is not interpreted as the percentage of misconceptions or even as the percentage of incorrect items because the item-level scoring structure is unavailable.

Score Interpretation

Achievement categories were used only as descriptive labels for the observed score distribution: very good (86-100), good (76-85), adequate (60-75), and low (<60). These categories provide a program-level summary and are not treated as validated proficiency cut scores. Interpretation therefore emphasizes the continuous-score statistics together with the category frequencies.

Statistical Procedures

Analysis was limited to statistics that directly answer the research questions: frequency and percentage by achievement category, mean, median, standard deviation, minimum, maximum, and a 95% confidence interval for the mean. The sample standard deviation used n-1 in the denominator, and the confidence interval used the t distribution because the population standard deviation was unknown. No inferential comparison among subgroups or science domains was attempted because the necessary variables were not available.

The analysis was intentionally parsimonious to avoid presenting multiple statistics that lead to the same substantive conclusion. The score distribution is therefore interpreted primarily through its central tendency, dispersion, observed range, confidence interval, and category composition.

Methodological clarification: a definitive diagnosis of misconception requires evidence beyond an aggregate score. In diagnostic assessment, an incorrect answer combined with a scientifically coherent but incorrect reason, and often a confidence judgment, provides stronger evidence of misconception than a single wrong response (Gurel et al., 2015; Treagust, 1988). The present study consequently uses the terms achievement, total score, and unearned-score proportion and does not report a misconception rate.

Result and Discussion

Participant Characteristics

The final analytic sample consisted of 80 prospective elementary teachers. All participants had completed or were taking the Basic Science Concepts course. De-identified student codes were used in the dataset.

Table 2. Descriptive Profile of Basic Science Achievement

Statistic	Value	Interpretation
Number of students	80	Observed cohort
Mean	77.88	Overall average score
Median	78.00	Central score
Standard deviation	3.23	Relatively narrow dispersion
Minimum-maximum	71-89	Observed score range
95% confidence interval	77.16-78.59	Estimated interval for the mean
Good category	65 (81.25%)	Largest achievement category
Unearned-score proportion	22.12 percentage points	Distance from the maximum score; not a misconception rate

The cohort obtained a mean score of 77.88 (SD = 3.23), a median of 78.00, and scores ranging from 71 to 89. The 95% confidence interval for the mean was 77.16-78.59. Based on the descriptive categories, 65 students (81.25%) were in the good category, 3 (3.75%) were very good, and 12 (15.00%) were adequate; no student was in the low category. Taken together, these values indicate a relatively concentrated distribution centered in the good range rather than a cohort containing sharply separated high- and low-performing groups.

The concentration of scores can be interpreted cautiously in light of the participants' shared curricular experience. All students were drawn from the same teacher-education program and had been exposed to the same or closely related course expectations, which may contribute to a relatively narrow score distribution. Prior research has shown that structured science coursework, hands-on inquiry, and reflective learning can improve prospective elementary teachers' conceptual understanding, confidence, and attitudes toward science (Avery & Meyer, 2012; Bleicher & Lindgren, 2005; Riegle-Crumb et al., 2015). The present cross-sectional data cannot demonstrate that those mechanisms produced the observed scores, but the literature makes shared instructional exposure a plausible contextual explanation for why the cohort performed within a relatively narrow band.

A second interpretation is that a homogeneous total-score distribution does not necessarily imply homogeneous conceptual understanding. Total scores can conceal compensatory patterns: a student may perform strongly on life-science items and weakly on force-and-motion items yet receive the same total score as a student with the opposite pattern. Domain-specific studies of prospective teachers have repeatedly shown that alternative conceptions can persist in particular topics such as astronomical phenomena, physical science, and biological processes (Atwood & Atwood, 1996; Brown & Schwartz, 2009; Burgoon et al., 2011). Therefore, the apparent uniformity in total scores should be understood as uniformity in overall test performance, not as evidence that all six science domains are equally well understood.

This distinction also explains why the study does not use the score distribution to identify difficult concepts. The available data do not show which items were missed, the reasons students selected particular answers, or the confidence with which answers were given. Diagnostic research demonstrates that these additional layers of evidence are necessary to distinguish a misconception from incomplete knowledge, guessing, or a careless error (Gurel et al., 2015; Nasir et al., 2024; Taşçı, 2024). Thus, the reviewer-relevant question 'what concepts are difficult and why?' remains empirically open and should be addressed in a follow-up study rather than inferred from the present total scores.

The mean of 77.88 leaves an average gap of 22.12 percentage points from the maximum possible score. This unearned-score proportion is educationally meaningful only at a general level: it signals that the cohort has room for additional learning even though most students are categorized as good. It should not be interpreted as a percentage of incorrect responses or misconceptions because the scoring architecture is unavailable. Possible contributors to the gap include incomplete knowledge, weak reasoning, item difficulty, unfamiliar contexts, or simple response error; the present dataset cannot determine their relative contribution.

For teacher education, this result supports a two-stage assessment strategy. Aggregate scores can be retained for routine program monitoring because they efficiently summarize cohort performance. When the goal shifts from monitoring to instructional diagnosis, however, the assessment should move to item-level evidence. Two-tier and multi-tier tests, constructed-response explanations, concept maps, and prediction-observation-explanation tasks can reveal the reasoning behind an answer and can therefore guide more targeted remediation (Duit & Treagust, 2003; Gurel et al., 2015; Treagust, 1988).

The interpretation is consistent with conceptual-change theory, which views learning as the restructuring of prior ideas rather than the simple accumulation of correct responses. Prospective teachers need opportunities to make their initial explanations visible, compare them with evidence, and revise them when existing conceptions are no longer adequate (Duit & Treagust, 2003; Posner et al., 1982). In this sense, a 'good' total score is a useful starting point but not an endpoint for science teacher preparation.

The need for explanatory depth is particularly important because prospective elementary teachers will later translate scientific knowledge into language, examples, demonstrations, and questions for children. Research on formative assessment and science practices suggests that teacher preparation should connect content learning with the capacity to elicit and respond to learners' ideas, rather than treating content knowledge and pedagogy as separate outcomes (Sabel et al., 2015; Seung et al., 2023).

The most important methodological implication of the findings is the separation between achievement measurement and misconception diagnosis. A conventional score answers 'how much of the assessment was earned,' whereas a diagnostic instrument is designed to answer 'what reasoning produced the response.' Treagust (1988) established the rationale for diagnostic tests that incorporate known alternative conceptions, and later reviews have shown the advantages and limitations of two-tier, three-tier, and four-tier formats (Gurel et al., 2015). Recent examples in light and primary-school biology similarly demonstrate that a defensible diagnostic instrument should report validity, reliability, and item characteristics rather than relying on total scores alone (Nasir et al., 2024; Taşçı, 2024).

For subsequent work with the Basic Science Concepts course, the original assessment should therefore be recovered and analyzed at item level before claims about domain difficulty or misconception prevalence are made. A stronger design would report the number and format of items, expert-review procedure, content-validity evidence, reliability, item difficulty, discrimination, and the distribution of response-reason combinations. If the original instrument cannot support such analysis, a new multi-tier diagnostic instrument should be developed and piloted across the six content domains.

The findings also have curricular implications. A cohort whose total scores are generally good does not necessarily require broad repetition of the entire course; instead, teacher educators need evidence that identifies where conceptual reinforcement is needed. Inquiry experiences and reflective practice can support science learning and teaching confidence, but their value is greatest when assessment reveals the ideas students are actually using (Menon & Azam, 2021; Riegle-Crumb et al., 2015; Strat et al., 2024).

A practical approach would combine a brief diagnostic pre-assessment with formative checks during each major science domain and a post-course assessment that includes both answers and reasoning. This would allow lecturers to separate lack of knowledge from stable alternative conceptions and to design remediation around specific causal explanations rather than around a generic score deficit.

Such a design would also strengthen the connection between science content knowledge and future classroom practice. Prospective teachers should be asked not only to select scientifically correct answers but also to explain concepts in age-appropriate language, anticipate common pupil ideas, design simple investigations, and evaluate evidence. These tasks move assessment closer to the knowledge demands of elementary science teaching.

Overall, the present results are best interpreted as a cohort-level baseline that identifies the need for more diagnostic evidence, not as a final account of students' conceptual structures. This bounded interpretation is more consistent with the available data and with current science-education measurement literature.

Based on the findings and their limitations, four actions are recommended. First, retain total-score reporting for routine cohort monitoring but report it together with dispersion and category frequencies.

Second, recover and psychometrically reanalyze the original instrument if item-level data remain available. Third, introduce multi-tier or reasoning-based diagnostic items across the six basic-science domains. Fourth, use the resulting diagnostic evidence to plan inquiry, conceptual-change, and formative-assessment activities targeted to specific patterns of student reasoning.

Limitations

This study has several limitations that directly constrain interpretation. First, the sample came from one teacher-education program and was not randomly selected from a broader population, so the findings are context-bound. Second, only total scores were available; item-level responses, domain subscores, reasoning, and confidence data were absent. Third, documentation required to independently evaluate the original assessment's psychometric quality—such as the number and format of items, expert-validation records, reliability, item difficulty, and discrimination—was not available for this analysis. The instrument is therefore treated as a course achievement measure rather than a validated diagnostic measure of concept mastery. Fourth, the cross-sectional design contains no pretest, comparison group, or longitudinal follow-up and cannot support causal claims about the course. The findings should consequently be read as a descriptive baseline and as a rationale for a more rigorous diagnostic follow-up. If the original instrument and item-level data can be recovered, reanalysis of their validity and reliability should precede stronger publication claims about specific conceptual mastery.

Conclusion

The observed cohort of 80 prospective elementary teachers achieved a mean Basic Science Concepts score of 77.88 (SD = 3.23), with most students classified in the good category and a relatively narrow score distribution. These results provide evidence of generally adequate-to-good overall course-test performance, but they do not identify which science domains are difficult or whether students hold specific misconceptions. The principal contribution of the study is therefore interpretive: favorable aggregate scores are useful for cohort monitoring, yet they should not be treated as a substitute for diagnostic evidence about students' reasoning. Future research should recover and analyze item-level data or employ validated multi-tier and reasoning-based instruments across the major basic-science domains so that conceptual reinforcement can be targeted on defensible evidence.

Acknowledgments
The authors thank the students, lecturers, and the Elementary School Teacher Education Program of Universitas Negeri Manado for supporting this research.

Author Contributions
Conceptualization, F.K. and Y.R.; methodology, F.K. and C.R.; validation, C.R. and Y.R.; formal analysis, F.K.; investigation, F.K. and C.R.; resources, Y.R.; data curation, F.K. and C.R.; writing-original draft preparation, F.K.; writing-review and editing, C.R. and Y.R.; visualization, F.K.; supervision, Y.R.; project administration, F.K. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare no conflict of interest.

Data Availability
Anonymized score data are available from the corresponding author upon reasonable request, subject to institutional and ethical requirements.

References

Addido, J., Burrows, A., & Slater, T. F. (2022). Addressing pre-service teachers' misconceptions and promoting conceptual understanding through the conceptual change model. Problems of Education in the 21st Century, 80(4), 477-498.

Atwood, V. A., & Atwood, R. K. (1995). Preservice elementary teachers' conceptions of what causes night and day. School Science and Mathematics, 95, 290-294.

Atwood, R. K., & Atwood, V. A. (1996). Preservice elementary teachers' conceptions of the causes of seasons. Journal of Research in Science Teaching, 33(5), 553-563. https://doi.org/10.1002/(SICI)1098-2736(199605)33:5<553::AID-TEA6>3.0.CO;2-Q

Avery, L. M., & Meyer, D. Z. (2012). Teaching science as science is practiced: Opportunities and limits for enhancing preservice elementary teachers' self-efficacy for science and science teaching. School Science and Mathematics, 112(7), 395-409. https://doi.org/10.1111/j.1949-8594.2012.00159.x

Bleicher, R. E., & Lindgren, J. (2005). Success in science learning and preservice science teaching self-efficacy. Journal of Science Teacher Education, 16(3), 205-225. https://doi.org/10.1007/s10972-005-4861-1

Brown, M. H., & Schwartz, R. S. (2009). Connecting photosynthesis and cellular respiration: Preservice teachers' conceptions. Journal of Research in

Science Teaching, 46(7), 791-812. https://doi.org/10.1002/tea.20287
Burgoon, J. N., Heddle, M. L., & Duran, E. (2011). Re-examining the similarities between teacher and student conceptions about physical science. Journal of Science Teacher Education, 22(2), 101-114. https://doi.org/10.1007/s10972-010-9196-x
d'Alessio, M. A. (2018). The effect of microteaching on science teaching self-efficacy beliefs in preservice elementary teachers. Journal of Science Teacher Education, 29(6), 441-467. https://doi.org/10.1080/1046560X.2018.1456883
Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. International Journal of Science Education, 25(6), 671-688. https://doi.org/10.1080/09500690305016
Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. Eurasia Journal of Mathematics, Science and Technology Education, 11(5), 989-1008. https://doi.org/10.12973/eurasia.2015.1369a
Howitt, C. (2007). Pre-service elementary teachers' perceptions of factors in an holistic methods course influencing their confidence in teaching science. Research in Science Education, 37, 41-58. https://doi.org/10.1007/s11165-006-9015-8
Koutsianou, A., & Emvalotis, A. (2022). Assessment of Greek pre-service primary teachers' efficacy beliefs in physics teaching. Journal of Science Teacher Education, 33(4), 389-409. https://doi.org/10.1080/1046560X.2021.2023959
Kruse, J., Henning, J., Wilcox, J., Carmen, K., Patel, N., & Seebach, C. (2021). Investigating the correlation between preservice elementary teachers' self-efficacy and science teaching practices. Journal of Science Teacher Education, 32(4), 469-479. https://doi.org/10.1080/1046560X.2020.1861767
Menon, D. (2020). Influence of the sources of science teaching self-efficacy in preservice elementary teachers' identity development. Journal of Science Teacher Education, 31(4), 460-481. https://doi.org/10.1080/1046560X.2020.1718863
Menon, D., & Azam, S. (2021). Investigating preservice teachers' science teaching self-efficacy: An analysis of reflective practices. International Journal of Science and Mathematics Education, 19, 1587-1607. https://doi.org/10.1007/s10763-020-10131-4
Menon, D., & Sadler, T. D. (2016). Preservice elementary teachers' science self-efficacy beliefs and science content knowledge. Journal of Science Teacher Education, 27(6), 649-673. https://doi.org/10.1007/s10972-016-9479-y
Moslemi, N., & Mousavi, A. (2019). A psychometric re-examination of the Science Teaching Efficacy and Beliefs Instrument in a Canadian context. Education Sciences, 9(1), 17. https://doi.org/10.3390/educsci9010017
Nasir, M., Cari, Sunarno, W., & Rahmawati, F. (2024). The development of a two-tier diagnostic test for student understanding of light. Journal of Turkish Science Education, 20(1), 150-172. https://doi.org/10.36681/tused.2023.009
Norris, C. M., Morris, J. E., & Lummis, G. W. (2018). Preservice teachers' self-efficacy to teach primary science based on 'science learner' typology. International Journal of Science Education, 40(18), 2292-2308. https://doi.org/10.1080/09500693.2018.1528645
Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66(2), 211-227. https://doi.org/10.1002/sce.3730660207
Riegle-Crumb, C., Morton, K., Moore, C., Chimonidou, A., Labrake, C., & Kopp, S. (2015). Do inquiring minds have positive attitudes? The science education of preservice elementary teachers. Science Education, 99(5), 819-836. https://doi.org/10.1002/sce.21177
Sabel, J. L., Forbes, C. T., & Zangori, L. (2015). Promoting prospective elementary teachers' learning to use formative assessment for life science instruction. Journal of Science Teacher Education, 26(4), 419-445. https://doi.org/10.1007/s10972-015-9431-6
Seung, E., Park, S., Kite, V., & Choi, A. (2023). Elementary preservice teachers' understandings and task values of the science practices advocated in the NGSS in the US. Education Sciences, 13(4), 371. https://doi.org/10.3390/educsci13040371
Strat, T. T. S., Henriksen, E. K., & Jegstad, K. M. (2024). Inquiry-based science education in science teacher education: A systematic review. Studies in Science Education, 60(2), 191-249. https://doi.org/10.1080/03057267.2023.2207148
Taşçı, G. (2024). Development of a two-tier diagnostic test to assess misconceptions about biology concepts at primary school. Science Insights Education Frontiers, 25(2), 4099-4122. https://doi.org/10.15354/sief.24.or652
Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. International Journal of Science Education, 10(2), 159-169. https://doi.org/10.1080/0950069880100204
Trundle, K. C., Atwood, R. K., & Christopher, J. E. (2002). Preservice elementary teachers' conceptions of moon phases before and after instruction. Journal of

Research in Science Teaching, 39(7), 633-658. https://doi.org/10.1002/tea.10039

Kikas, E. (2004). Teachers' conceptions and misconceptions concerning three natural phenomena. Journal of Research in Science Teaching, 41(5), 432-448. https://doi.org/10.1002/tea.20012

Kanli, U. (2014). A study on identifying the misconceptions of pre-service and in-service teachers about basic astronomy concepts. Eurasia Journal of Mathematics, Science and Technology Education, 10(5), 471-479. https://doi.org/10.12973/eurasia.2014.1120a

Starakis, I., & Halkia, K. (2014). Addressing K-5 students' and pre-service elementary teachers' conceptions of seasonal change. Physics Education, 49(2), 231-239. https://doi.org/10.1088/0031-9120/49/2/231

Trumper, R. (2006). Teaching future teachers basic astronomy concepts—seasonal changes—at a time of reform in science education. Journal of Research in Science Teaching, 43(9), 879-906. https://doi.org/10.1002/tea.20138

Narjaikaew, P. (2013). Alternative conceptions of primary school teachers of science about force and motion. Procedia - Social and Behavioral Sciences, 88, 250-257. https://doi.org/10.1016/j.sbspro.2013.08.503