

Written Essay Assessment of Philosophical Reasoning in Game-Based Science Teacher Education: Ontological, Epistemological, and Axiological Profiles of Pre-Service Teachers

Rifda Nur Hikmahwati Arif*, Banu Kisworo, Yulianti Yusal

Makassar State University, Indonesia

*Corresponding author: rifdanha@unm.ac.id

Abstract: Philosophy of Science is a key component of science teacher education because it helps prospective teachers understand the nature of scientific knowledge, its justification, and its educational value. This descriptive qualitative study analyzed essay responses from 78 pre-service science teachers at Makassar State University who participated in a multiplatform game-based Philosophy of Science course. Four essay questions assessed understanding of ontology, epistemology, axiology, and the integration of these three dimensions within junior secondary science education. Data were examined through qualitative content analysis supported by codebook development, repeated reading, peer review, negative-case analysis, and descriptive quantification. Results indicated strong understanding of ontological concepts (95.0% valid responses) and axiological reasoning (96.3%), while epistemological understanding was less developed (59.6%). Integrating the three philosophical dimensions proved most challenging, with only 53.8% valid responses and 40.5% demonstrating complete synthesis. Students primarily recommended debate-based games (31.3%), reflecting an emphasis on critical thinking. These findings provide a framework for assessing philosophical reasoning through essays in game-based science teacher education.

Keywords: philosophy of science, pre-service science teachers, ontology, epistemology, axiology, game-based learning, qualitative content analysis, written reasoning assessment

Abstrak: Filsafat sains merupakan komponen penting dalam pendidikan calon guru sains karena membantu mereka memahami hakikat pengetahuan ilmiah, dasar pembenarannya, serta nilai dan tujuan pendidikan yang dikandungnya. Penelitian kualitatif deskriptif ini menganalisis respons esai dari 78 mahasiswa calon guru sains di Universitas Negeri Makassar yang mengikuti perkuliahan Filsafat sains berbasis permainan multiplatform. Empat pertanyaan esai digunakan untuk menilai pemahaman tentang ontologi, epistemologi, aksiologi, serta integrasi ketiga dimensi tersebut dalam konteks pembelajaran IPA SMP. Data dianalisis menggunakan analisis isi kualitatif yang didukung oleh pengembangan kode, pembacaan berulang, diskusi sejawat, pemeriksaan kasus negatif, dan kuantifikasi deskriptif. Hasil penelitian menunjukkan pemahaman yang kuat pada aspek ontologi (95,0% respons valid) dan penalaran aksiologis (96,3%), sedangkan pemahaman epistemologis masih kurang berkembang (59,6%). Integrasi ketiga dimensi filsafat menjadi aspek yang paling menantang, dengan hanya 53,8% respons valid dan 40,5% yang menunjukkan sintesis lengkap. Mahasiswa paling banyak merekomendasikan permainan berbasis debat (31,3%), yang mencerminkan penekanan pada pengembangan berpikir kritis. Temuan ini memberikan kerangka untuk menilai penalaran filosofis melalui esai dalam pendidikan calon guru sains berbasis permainan.

Kata Kunci: filsafat sains, calon guru sains, ontologi, epistemologi, aksiologi, pembelajaran berbasis permainan, analisis konten kualitatif, penilaian penalaran tertulis

INTRODUCTION

Philosophy of Science holds an important position in science teacher education because it enables prospective teachers to understand science as more than a collection of facts, formulae,

and procedures. Through philosophical reflection, pre-service science teachers learn to examine the nature of scientific objects, the justification of scientific claims, and the ethical purposes of science education. These dimensions are addressed through three pillars: ontology (what exists and what science investigates), epistemology (how scientific knowledge is acquired, justified, and revised), and axiology (the values, ethics, and purposes of scientific knowledge in human life and education) (Abd-El-Khalick, 2005; Erduran & Dagher, 2014; Lederman et al., 2002).

For pre-service science teachers, these pillars directly shape pedagogical decisions. A teacher's ontological stance influences how scientific entities are presented as objects of inquiry; epistemological understanding shapes how evidence, verification, and the tentative nature of knowledge are taught; axiological understanding helps teachers connect science to ethical responsibility and meaningful action (Allchin, 2011; Matthews, 2015; McComas, 2020; Taber, 2013).

Despite their importance, philosophical understandings are frequently assessed through questionnaires, multiple-choice tests, or Likert-scale instruments that provide limited evidence of reasoning quality. Written essays offer richer evidence because they allow students to construct explanations, use examples, expose misconceptions, and demonstrate whether they can connect philosophical concepts to science teaching practice (Zohar & Dori, 2003; Erduran et al., 2022). Essays are not merely assessment products they are windows into students' developing philosophical reasoning.

Game-based learning (GBL) has been proposed as an approach to making abstract content more concrete, interactive, and emotionally engaging (Clark et al., 2016; Plass et al., 2015; Wouters et al., 2013). In Philosophy of Science, games such as concept debates, ontology puzzles, philosophical role-plays, and theory-revision simulations can help students experience philosophical tension rather than merely memorise definitions. However, high engagement in GBL does not automatically guarantee deep philosophical synthesis students may enjoy the games yet still struggle to articulate the tentative nature of science, apply axiology as a philosophical commitment, or integrate all three pillars coherently.

The use of multiplatform GBL in Philosophy of Science for pre-service science teachers remains underexplored, and little is known about how students express philosophical understanding after such learning. This study addresses four research questions: (1) How completely do students respond to essay prompts on each philosophical pillar? (2) What dominant themes emerge across each set of responses? (3) Which aspects of philosophical reasoning are strong, partial, or underdeveloped? and (4) What implications can be drawn for improving the design of game-based Philosophy of Science learning?

The novelty of this research lies in combining a multiplatform GBL context with essay-based assessment of three-pillar philosophical reasoning an approach not previously documented in science teacher education in Indonesia. This focus is further justified by prior research on the nature of science, which has demonstrated that improvements in pre-service teachers' NOS conceptions do not automatically translate into instructional planning or classroom action (Bell et al., 2000). Science educators have argued that learners should be exposed to explicit ideas about evidence, explanation, creativity, uncertainty, and social responsibility in science (Osborne et al., 2003), whilst the implementation of history and philosophy of science remains difficult because teachers frequently face curricular, epistemological, and practical constraints (Höttecke & Silva, 2011). Accordingly, the essay-based assessment in this study is positioned not merely as a measure of conceptual recall, but as a means of examining how pre-service teachers argue about science, pedagogy, and contextualised decision-making (Khishfe, 2012).

METHODS

Research Design

This study employed a descriptive qualitative design supported by simple quantitative summaries (Creswell & Poth, 2018; Merriam & Tisdell, 2016). Qualitative content analysis was the primary analytical approach, combining inductive coding with concept-sensitive interpretation

grounded in the three philosophical pillars (Hsieh & Shannon, 2005; Krippendorff, 2019; Mayring, 2014). The thematic coding procedure followed the flexible logic of thematic analysis, suited to identifying patterned meanings across qualitative responses whilst allowing interpretation to remain theoretically grounded (Braun & Clarke, 2006). Coding decisions were further guided by a qualitative coding manual emphasising systematic code development, category refinement, and transparent analytical documentation (Saldaña, 2021).

Participants and Learning Context

Participants comprised 78 pre-service science teachers (Cohort 2024) from the Department of Science Education, Universitas Negeri Makassar, enrolled in three parallel class sections during the 2025/2026 academic year: Class A (n=31), Class B (n=16), and Class C (n=31). The course employed a multiplatform GBL design integrating offline games (Concept Puzzle O-E-A, Socratic Discussion Cards, Ontology Word Battle, Philosopher's Journey Board Game), psychomotor activities (Standing Debate, Integrated Concept Embodiment, Gallery Walk, Scientific Experiment Drama), and digital platforms (Quizizz, Wordwall, Kahoot!, Padlet, digital escape rooms). All participants completed game-based activities prior to essay assessment.

The use of varied game formats was grounded in principles of game-based learning environments, particularly the view that well-designed games can create problem-solving spaces with feedback, challenge, and progressive participation (Gee, 2003). Empirical research on challenging games further indicates that engagement, immersion, and flow can support learning when game challenges are meaningfully aligned with instructional objectives (Hamari et al., 2016). The digital escape room component was included because educational escape rooms have been reported to foster collaboration and engagement when puzzles are explicitly linked to desired learning outcomes (Veldkamp et al., 2020).

Essay Instrument

Four essay prompts were administered following completion of game-based activities, progressing from component-level understanding to integrative synthesis (Table 1).

Table 1. Essay prompts by philosophical pillar

Essay	Pillar	Prompt (administered in Indonesian)
1	Ontology	Explain the difference between realist and idealist views in the ontology of science. How do these two views influence the way a junior secondary science teacher perceives the objects being taught? Provide an example from junior secondary science content.
2	Epistemology	How do scientists prove the truth about photosynthesis? Connect this to epistemological concepts in science: sources of knowledge, verification, and the tentative nature of science.
3	Axiology	A junior secondary science teacher is teaching a unit on waste and environmental pollution. Explain how the axiological dimension of science can guide this teacher in designing meaningful, character-building learning.
4	Integration	Synthesise the three pillars of philosophy of science ontology, epistemology, and axiology using one junior secondary science topic of your choice. Explain how all three complement each other in building a complete understanding of science.

Note. Prompts were delivered in the original Indonesian language as administered.

Data Analysis

Analysis proceeded in five stages: (1) screening valid from invalid responses (blank, dash, slash entries); (2) calculating response completeness per item and class; (3) measuring response length by character count; (4) inductively coding valid responses and grouping them into themes for Essays 1–3, and classifying integration quality levels for Essay 4; and (5) interpreting themes against philosophical concepts with representative verbatim quotations selected per theme. Minor spelling errors in quotations were corrected for readability without altering meaning.

Trustworthiness

Credibility was strengthened through repeated reading, cross-prompt code comparison, and representative quotations (Lincoln & Guba, 1985; Nowell et al., 2017). Dependability was supported through a coding matrix. Confirmability was strengthened through peer debriefing two colleagues independently reviewed a 20% random sample of coded responses, reaching agreement prior to final coding and through negative-case checking for partial or ambiguous responses. Transferability was supported through detailed description of participants, course design, prompts, and analytical categories.

Ethical Considerations

This study used naturally occurring course assessment responses for academic reflection. Student identities were not reported; quotations were anonymised using only class section labels. Analysis focused on aggregate patterns rather than individual performance evaluation.

RESULTS AND DISCUSSION

Response Completeness and Length

Table 2 presents response completeness and mean response length. Essay 1 (Ontology) yielded the highest valid response rate (76.9%), whilst Essay 4 (Integration) yielded the lowest (53.8%), indicating that synthesis tasks are more cognitively demanding. Notably, valid Essay 4 responses were the longest ($M = 1,446$ characters), suggesting that students who attempted integration invested greater elaborative effort.

Table 2. Response completeness and mean length by essay item and class section

Essay Item	Class A (n=31)	Class B (n=16)	Class C (n=31)	Total Valid n (%)	M Characters
1. Ontology	25 (80.6%)	14 (87.5%)	21 (67.7%)	60 (76.9%)	1,178
2. Epistemology	22 (71.0%)	12 (75.0%)	18 (58.1%)	52 (66.7%)	1,288
3. Axiology	22 (71.0%)	13 (81.3%)	19 (61.3%)	54 (69.2%)	1,143
4. Integration	19 (61.3%)	9 (56.3%)	14 (45.2%)	42 (53.8%)	1,446

Note. *M chars = mean character count of valid responses only. Invalid responses were excluded from qualitative analysis.*

The progressive decline from Essay 1 to Essay 4 mirrors the Cognitive Engagement pattern observed in student engagement surveys from the same course ($M = 4.099$ for Cognitive vs. $M = 4.442$ for Behavioural), providing convergent evidence that game-based learning produces stronger participation and affect than philosophical synthesis.

Essay 1: Ontological Understanding Realism and Idealism

Analysis of 60 valid responses yielded five themes (Table 3). The nature/independence dichotomy (T1) appeared in 95.0% of responses, indicating strong definitional mastery of the core ontological contrast. Most students explained that realism affirms the mind-independent existence of scientific objects, whilst idealism makes them dependent on human thought or perception. The

majority also connected this to science teaching (T2; 76.7%) and cited specific junior secondary science examples (T3; 73.3%).

Table 3. Thematic analysis of Essay 1 responses: Ontology (n = 60)

Theme	Description		f	%
T1	Nature/independence dichotomy	Realism = mind-independent scientific objects; idealism = thought/perception-dependent reality	57	95.0%
T2	Pedagogical application	Linking ontological stance to how junior secondary science teachers explain objects, facts, models, and concepts	46	76.7%
T3	Concrete junior secondary science examples	Atoms, cells, gravity, photosynthesis, planetary systems, or ecosystems mentioned	44	73.3%
T4	Complementarity argument	Effective science teaching can integrate both realist and idealist orientations	18	30.0%
T5	Partial/blurred understanding	Conflating realism-idealism with empiricism-rationalism; lack of pedagogical application	9	15.0%

Note. Themes are non-exclusive; one response may be categorised under multiple themes.

"Realism holds that scientific objects such as cells, forces, atoms, or energy genuinely exist in reality, independent of the human mind. This means that even when they are not directly observable, these objects still exist and can be studied through scientific methods." (Participant, Class C)

The complementarity argument (T4; 30.0%) represents the philosophically most sophisticated pattern, acknowledging that realism and idealism are not mutually exclusive but can be productively combined in science teaching. This aligns with the Family Resemblance Approach to the nature of science (Erduran & Dagher, 2014). The 15.0% rate of partial understanding (T5) indicates that a minority of respondents still require explicit guidance in distinguishing ontological positions from epistemological ones.

Essay 2: Epistemological Reasoning The Photosynthesis Case

Among 52 valid responses, strong engagement with empirical experimentation and verification was evident, yet tentative thinking was less developed (Table 4). Whilst 94.2% addressed experimentation and 73.1% addressed verification, only 59.6% explicitly discussed the revision of scientific knowledge the most philosophically sophisticated epistemological dimension.

Table 4. Thematic analysis of Essay 2 responses: Epistemology (n = 52)

Theme	Description		f	%
T1	Empirical experimentation	Knowledge about photosynthesis obtained through systematic observation and experimentation	49	94.2%

T2	Verification	Claims verified through repeated testing, controlled variables, and replication	38	73.1%
T3	Tentative nature	Knowledge about photosynthesis is open to revision with new evidence or methods	31	59.6%
T4	Historical experimentation	Scientists/experiments mentioned: Ingenhousz, Sachs, van Helmont, Ruben-Kamen isotope tracing	22	42.3%
T5	Full integration	All three epistemological dimensions discussed coherently in a single response	19	36.5%
T6	Surface-level only	General mention of experiments without explaining verification or tentative nature	14	26.9%

Note. T6 represents responses that remained general despite mentioning experiments.

"Scientists proved photosynthesis through the epistemology of science: the source of knowledge derives from experiments such as those conducted by Ingenhousz and Sachs. Verification was done through repeated testing and variable control. However, the science is tentative it was once thought that O₂ came from CO₂, but the Ruben-Kamen experiment with ¹⁸O proved that O₂ actually comes from H₂O." (Participant, Class C)

The lower frequency of tentative reasoning (59.6%) carries important diagnostic significance. Tentativeness is one of the most philosophically significant yet educationally challenging dimensions of NOS (Lederman et al., 2002; McComas, 2020). Science teachers who do not understand tentativeness risk presenting science as dogmatic fact or mere opinion both epistemologically distorted. Future game designs should incorporate history-of-science scenarios that explicitly portray how understanding of photosynthesis was revised, such as a theory-revision game using Ruben and Kamen's isotope tracing evidence.

Essay 3: Axiological Reasoning Environmental Science Teaching

Essay 3 yielded the most contextually consistent and practice-grounded responses (Table 5). Nearly all valid responses (96.3%) connected science teaching to values and character formation, and 72.2% proposed concrete, action-oriented learning activities. This practical strength is promising for science teacher education.

Table 5. Thematic analysis of Essay 3 responses: Axiology (n = 54)

Theme	Description		f	%
T1	Values and character formation	Teaching directed towards responsibility, empathy, and moral awareness	52	96.3%
T2	Real-world contextualisation	Connecting pollution to daily life, local rivers, school waste, or plastic pollution	46	85.2%
T3	Action-oriented design	Waste sorting projects, composting, recycling, upcycling, eco-enzyme activities proposed	39	72.2%
T4	Beyond knowledge towards wisdom	Distinguishing between knowing about pollution and acting wisely based on science	28	51.9%

T5	Explicit ethical/philosophical framing	Terms used: intergenerational responsibility, environmental literacy, science as value-laden	18	33.3%
----	--	--	----	-------

"The axiological dimension of science focuses on the values, benefits, and ethics of knowledge. When a junior secondary science teacher teaches about waste and environmental pollution, axiology guides the learning so that it does not stop at 'knowing the definition and types of waste,' but reaches 'being willing and able to act.'" (Participant, Class C)

However, fewer responses (T5; 33.3%) addressed axiology as a philosophical claim that science is never value-neutral and carries an inherent ethical dimension regardless of its application. Students need to understand axiology not merely as a moral message added onto science lessons, but as a meta-scientific commitment about the relationship between knowledge and values (McComas, 2020; Allchin, 2011). Ethical dilemma games and scientific trial simulations may be particularly effective in developing this deeper dimension.

Essay 4: Integrative Philosophical Synthesis

Essay 4 yielded the lowest response rate (53.8%) but the longest valid responses (M = 1,446 characters). Table 6 shows that photosynthesis (33.3%) and pollution (21.4%) dominated topic selection the same contexts used in Essays 2 and 3 suggesting transfer from familiar rather than novel contexts. Integration quality is presented in Table 7.

Table 6. Junior secondary science topics selected for philosophical integration (n = 42)

Topic	Frequency	%
Photosynthesis	14	33.3%
Pollution and Environmental Waste	9	21.4%
Ecosystems	3	7.1%
Solar System	3	7.1%
Human Digestive / Excretory / Circulatory Systems	4	9.5%
States of Matter / Pressure / Light	3	7.1%
Other topics (Global Warming, Origins of Life, etc.)	6	14.3%

Table 7. Levels of philosophical integration in Essay 4 responses (n = 42)

Level	Characteristics	f	%
Level 3 Full integration	All three pillars explicitly discussed and interrelated; explains why they complement each other and what would be lost if any one were absent	17	40.5%
Level 2 Partial integration	All three pillars mentioned but discussed separately; limited argument on their interdependence	18	42.9%
Level 1 Minimal integration	One or two pillars discussed substantively; third pillar addressed superficially or absent	7	16.7%

"Ontology provides the subject matter: What do we study? Epistemology provides the method: How do we know it is true? Axiology provides the value direction: Why do we study it? With only ontology, students merely memorise definitions. With only

epistemology, students can test but do not care. With only axiology, students have values but lack knowledge." (Participant, Class A)

The 'limping' metaphor—the idea that removing any one pillar renders philosophical understanding unbalanced—emerged spontaneously across multiple Level 3 responses. This is a genuine philosophical insight consistent with the Family Resemblance Approach (Erduran & Dagher, 2014). Only 40.5% of valid respondents achieved full integration, indicating that game-based learning is effective in building component-level philosophical understanding, but synthesis-level reasoning requires more intentional scaffolding.

Student-Generated Game Suggestions

A total of 67 valid open-ended game suggestions were collected. Table 8 summarises their thematic categories. Debate-based games dominated (31.3%), followed by quiz/interactive games (26.9%) and guessing/problem-solving activities (25.4%).

Table 8. Thematic categories of student-generated game suggestions (n = 67)

Category	Examples from Responses	n	%
Debate-based games	Realism vs. Idealism vs. Constructivism Role-Play; Rotating Debate; pro-con contemporary science issues (genetic engineering, climate change)	21	31.3%
Quiz and interactive games	OEA Philosophy interactive quiz, trivia, true-false game, team question competition	18	26.9%
Guessing and problem-solving	Picture guessing, scientist guessing, Black Box Mystery simulation, philosophy puzzles	17	25.4%
Case study activities	History-of-science case studies, "What If" history of science, scientist ethical dilemma analysis	10	14.9%
Simulations and role-plays	Peer-review simulation, scientific journal simulation, Karl Popper/Thomas Kuhn role-play	10	14.9%
Escape rooms and mysteries	OEA digital escape room, closed-concept philosophy puzzles	4	6.0%

Note. Categories are non-exclusive.

"The 'Black Box' simulation: students determine the contents of a sealed box without opening it, using indirect experimentation—excellent for simulating the epistemology of science and the limits of observation." (Participant, Class C)

The sophistication of these suggestions—Black Box Mystery, three-position ontological debate, and science ethics dilemmas—provides evidence of metacognitive awareness: students were not only engaging with philosophical content but also thinking about how it is best learned through games (Deterding et al., 2011; Whitton, 2018).

DISCUSSION

Written Essays as an Assessment Tool for Philosophical Reasoning

This study confirms that written essays reveal important dimensions of philosophical reasoning that cannot be captured by closed-ended instruments. Whilst Likert or multiple-choice items efficiently identify broad patterns, essays expose the coherence, contextualisation, and integrative quality of student thinking. Students who scored highly on engagement measures were not always those who demonstrated the deepest philosophical integration in essays—reinforcing the distinction between motivational engagement and cognitive depth in GBL contexts (Wouters et al., 2013; Erhel & Jamet, 2013).

Patterns Across the Three Philosophical Pillars

Ontological understanding was strongest, likely because debate activities and concept games (Standing Debate, Ontology Word Battle) required students to actively take positions and defend assumptions precisely what philosophical argument demands (Lipman, 2003). Axiological reasoning was practically robust yet theoretically shallow: students understood axiology as a useful tool for teaching environmental values, but few engaged with its meta-scientific claim that science is never value-neutral (McComas, 2020). Epistemological tentativeness was the weakest dimension, suggesting that whilst game-based activities effectively conveyed empirical and verification dimensions, the historically contingent and revisable nature of scientific knowledge requires more deliberate game design such as theory-revision scenarios depicting how understanding of photosynthesis changed through Ruben and Kamen's isotope evidence.

Implications for Game-Based Philosophy of Science Learning Design

Four design implications emerge. First, debate and argumentation games are closely aligned with the demands of philosophical reasoning and should be systematically prioritised. Second, epistemological tentativeness requires dedicated history-of-science game designs. Third, axiology benefits from ethical dilemma games rather than only environmental action projects. Fourth, philosophical integration cannot be achieved through a single synthesis task it requires repeated three-lens analysis embedded throughout the course using different science topics each time to prevent over-reliance on familiar contexts.

The role of post-game debriefing sessions is also critical. Game activities must be paired with structured reflection that explicitly connects game experiences to philosophical concepts (Crookall, 2010). Without debriefing, students may engage enthusiastically in games yet miss the philosophical meaning the games were intended to convey.

Limitations

Several limitations warrant acknowledgement. First, the descriptive design without a control group or pre-post comparison prevents causal claims. Second, a single-institution sample limits transferability. Third, the analysis focuses on written responses and cannot capture oral reasoning or reasoning over time. Fourth, although peer debriefing on themes was conducted, inter-rater reliability coefficients (e.g., Krippendorff's alpha) were not formally calculated a limitation that should be addressed in future research. Fifth, the low response rate for Essay 4 (53.8%) may introduce selection bias if non-respondents systematically differ from respondents in philosophical understanding.

CONCLUSION

This descriptive qualitative study analysed the written philosophical reasoning of 78 pre-service science teachers enrolled in a multiplatform game-based Philosophy of Science course at Universitas Negeri Makassar. Four main findings emerged: (1) ontological understanding was strongest, with 95.0% of valid responses correctly characterising the realism-idealism distinction; (2) axiological reasoning was practice-grounded but philosophically shallow, particularly regarding science as an inherently value-laden domain; (3) epistemological tentativeness was the least developed dimension (59.6%), despite strong empirical and verification reasoning; and (4) philosophical integration was the most challenging aspect, with only 53.8% valid response rate and 40.5% achieving full three-pillar synthesis.

These findings provide a descriptive framework for using written essays to assess philosophical reasoning quality in GBL-based Philosophy of Science courses, and identify specific scaffolding targets: tentativeness games, ethical dilemma activities, and three-lens integration tasks embedded throughout rather than only at the end of the course. Future research should formally calculate inter-rater reliability coefficients, develop and validate philosophical reasoning rubrics, and extend to longitudinal or quasi-experimental designs.

ACKNOWLEDGMENTS

The authors wish to thank the pre-service science teachers of Cohort 2024 in the Department of Science Education, Universitas Negeri Makassar, for their participation and written responses.

REFERENCES

- Abd-El-Khalick, F. (2005). Developing deeper understandings of nature of science: The impact of a philosophy of science course on preservice science teachers' views and instructional planning. *International Journal of Science Education*, 27(1), 15–42. <https://doi.org/10.1080/0950069042000180910>
- Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education*, 95(3), 518–542. <https://doi.org/10.1002/sce.20432>
- Bell, R. L., Lederman, N. G., & Abd-El-Khalick, F. (2000). Developing and acting upon one's conception of the nature of science: A follow-up study. *Journal of Research in Science Teaching*, 37(6), 563–581.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.
- Crookall, D. (2010). Serious games, debriefing, and simulation/gaming as a discipline. *Simulation & Gaming*, 41(6), 898–920. <https://doi.org/10.1177/1046878110390784>
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. *Proceedings of the 15th International Academic MindTrek Conference*, 9–15. <https://doi.org/10.1145/2181037.2181040>
- Erduran, S., & Dagher, Z. R. (2014). *Reconceptualizing the nature of science for science education: Scientific knowledge, practices and other family categories*. Springer.
- Erduran, S., Kaya, E., & Dagher, Z. R. (2022). From the margins to the mainstream: The contested positioning of nature of science in science education. *Science & Education*, 31(6), 1723–1742. <https://doi.org/10.1007/s11191-022-00390-0>
- Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, 67, 156–167. <https://doi.org/10.1016/j.compedu.2013.02.019>
- Gee, J. P. (2003). What video games have to teach us about learning and literacy. Palgrave Macmillan.
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54, 170–179. <https://doi.org/10.1016/j.chb.2015.07.045>
- Höttecke, D., & Silva, C. C. (2011). Why implementing history and philosophy in school science education is a challenge: An analysis of obstacles. *Science & Education*, 20(3-4), 293–316. <https://doi.org/10.1007/s11191-010-9285-4>
- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Khishfe, R. (2012). Nature of science and decision-making. *International Journal of Science Education*, 34(1), 67–100. <https://doi.org/10.1080/09500693.2011.559490>
- Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE.

- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39(6), 497–521. <https://doi.org/10.1002/tea.10034>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- Lipman, M. (2003). *Thinking in education* (2nd ed.). Cambridge University Press.
- Matthews, M. R. (2015). *Science teaching: The contribution of history and philosophy of science* (20th anniversary revised and expanded edition). Routledge.
- Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Beltz.
- McComas, W. F. (Ed.). (2020). *Nature of science in science instruction: Rationales and strategies*. Springer.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1–13. <https://doi.org/10.1177/1609406917733847>
- Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What 'ideas-about-science' should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, 40(7), 692–720. <https://doi.org/10.1002/tea.10105>
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE.
- Taber, K. S. (2013). Conceptual frameworks, metaphysical commitments and worldviews: The challenge of reflecting on the relationship between science and religion in science education. In N. Mansour & R. Wegerif (Eds.), *Science education for diversity: Theory and practice* (pp. 151–177). Springer. https://doi.org/10.1007/978-94-007-4563-6_8
- Veldkamp, A., van de Grint, L., Knippels, M. C. P. J., & van Joolingen, W. R. (2020). Escape education: A systematic review on escape rooms in education. *Educational Research Review*, 31, 100364. <https://doi.org/10.1016/j.edurev.2020.100364>
- Whitton, N. (2018). Playful learning: Tools, techniques and tactics. *Research in Learning Technology*, 26, 2035. <https://doi.org/10.25304/rlt.v26.2035>
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>
- Zohar, A., & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? *Journal of the Learning Sciences*, 12(2), 145–181. https://doi.org/10.1207/S15327809JLS1202_1