INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 14, Issue 4, (2025) E-ISSN: 2226-6348

What Works Outdoors? A Systematic Literature Review of Nature-Based Instruction in Science Education

Siti Nur Diyana Mahmud* & Muhammad Hanif Sainorudin

Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Corresponding Author Email: diyana@ukm.edu.my

DOI Link: http://dx.doi.org/10.6007/IJARPED/v14-i4/26729

Published Online: 24 October 2025

Abstract

This systematic literature review examines methodological trends in nature-based instruction (NBI) in science education between 2021 and 2025, focusing on formal and school-linked learning settings. Based on searches in Scopus and Web of Science, 138 empirical studies were screened, with 58 meeting quality criteria using the Mixed Methods Appraisal Tool (MMAT). Most studies used qualitative or mixed-methods designs, drawing primarily on observations, field notes, and concept inventories. Analytic approaches leaned heavily on thematic and content analysis. Pedagogically, outdoor inquiry and place-based education were the most common strategies, supported by school gardens and, to a lesser extent, citizen science and TEK-informed programs. Outcomes focused mainly on affective domains (e.g., identity, interest, self-efficacy, nature connectedness) and cognitive gains (e.g., achievement, knowledge), with fewer studies addressing scientific practices or behavioural change. Most interventions occurred in nearby outdoor sites and were delivered through formal education. While the field shows strength in ecological validity and instructional design, recurring challenges include limited use of causal designs, short follow-up periods, and inconsistent implementation reporting. The review recommends building a shared measurement framework, expanding equity-focused research, and integrating AI tools responsibly to support future growth in outdoor science education.

Keywords: Nature-based Instruction, Outdoor Learning, Place-Based Education, School Gardens, Citizen Science, Science Education, Inquiry-Based Learning, Cognitive and Affective Outcomes, Data Literacy, Mixed-Methods Research, PRISMA, MMAT

Introduction

Nature-based instruction (NBI) is drawing renewed attention as schools adapt after the pandemic, climate and sustainability move up policy agendas, equity debates foreground Traditional Ecological Knowledge (TEK), and educators begin experimenting with AI and learning analytics in outdoor learning. These shifts call for a synthesis of what we know, rather than more isolated examples. Yet the field remains methodologically fragmented. Research designs, measurement tools, reporting practices, and implementation details are inconsistent, making comparison and cumulative evidence difficult.

The years since 2020 have deepened the relevance of nature-based instruction (NBI). In response to the disruptions caused by the pandemic, educational systems around the world began to reconsider where, how, and with what tools students engage in learning. Outdoor environments and local micro-sites offered safer, more adaptable, and engaging alternatives that made it possible to continue instruction even as public health guidelines shifted. At the same time, the escalating climate crisis pushed issues like sustainability, biodiversity loss, and socio-ecological resilience to the forefront of both policy and educational discourse. These shifts happened alongside rapid advancements in artificial intelligence (AI) and data technologies, which opened up new avenues for designing learning experiences, delivering feedback, and analyzing student work, while also raising important concerns around privacy, equity, and the role of learners in shaping knowledge. Collectively, these forces have reshaped not only what science is taught, but also how research into science education is carried out: where studies take place, which designs are viable, what counts as valid evidence, and which outcomes are given priority.

Historically, research into nature-based instruction (NBI) has offered vivid accounts of student engagement, interest, and emotional connection to the natural world, alongside encouraging evidence of growth in conceptual understanding, inquiry skills, and environmental attitudes. However, the field's methodological diversity has made it difficult to build a cohesive and cumulative body of knowledge. Much of the work has relied on practitioner-led case studies, small pilot projects, and quasi-experimental designs using pre- and post-assessments, often with tools tailored to specific contexts and limited long-term tracking. These features aren't accidental or insignificant, they reflect the unique demands of outdoor learning environments, such as dealing with weather, terrain, safety, and access. They also reflect the ethical complexities of working with communities and local knowledge holders, which require trust, reciprocity, and respect for data sovereignty. Additionally, the constraints of formal schooling are schedules, resources, and the need to align with assessment frameworks to further shape what is feasible. These challenges raise a key methodological question: How can the field maintain the authenticity of nature-based learning while introducing the rigor needed to build a robust and transferable knowledge base?

There is a pressing need for a methodologically focused synthesis of NBI research in today's context. Since 2021, researchers have been experimenting with blended approaches to measurement, combining tools such as pre- and post-tests, Likert-scale surveys, interviews, observations, and, in some cases, structured lesson observation protocols (e.g., modified versions of RTOP or COPUS adapted for outdoor learning environments). A smaller portion of studies has begun incorporating logs, sensors, geotagged photographs, or low-stakes analytics to track student activity. In parallel, research on teacher education has expanded its focus to include both preservice and inservice teachers' self-efficacy, their ability to design effective instruction, and their fidelity in implementing NBI programs. Citizen-science models have brought additional attention to issues such as data quality, participation trends, and environmental stewardship behaviours. However, the use of rigorous causal research designs like randomized controlled trials, regression discontinuity, or stepped-wedge rollouts remains uncommon, largely due to practical and ethical challenges, especially when it's not acceptable to withhold valuable outdoor learning experiences from some students. Long-term studies that assess whether changes in understanding or attitudes endure over time are also rare, with most research reporting only on outcomes measured immediately after the intervention.

Additionally, inconsistent reporting on how faithfully programs are implemented and the quality of instruction makes it difficult to interpret results and replicate studies.

Emphasizing the importance of place and community in nature-based instruction (NBI) also brings forward critical issues of equity and knowledge representation. In many school systems, historically underserved communities face significant barriers which is limited access to safe and well-maintained green spaces, fewer resources for transportation and materials, and restricted scheduling flexibility. This makes on-campus micro-sites and nearby urban greenspaces especially valuable, but also methodologically sensitive. Researchers must be careful not to frame these contexts or the learners within them from a deficit perspective, while still acknowledging the structural constraints at play. Integrating Traditional Ecological Knowledge (TEK) and culturally sustaining pedagogies demands thoughtful methodological approaches such as collaborative design with community partners, ethical review processes that reflect local priorities, and data practices that respect community ownership and sovereignty. As studies increasingly span different cultural and regional settings, it's important to balance measurement consistency with cultural sensitivity and language nuance. A methodologically driven review can therefore reveal not just technical considerations like instrument selection, but also the ethical foundations that shape research design.

At the same time, the rise of artificial intelligence (AI) in education has opened a new frontier for research in nature-based instruction (NBI). Some recent studies have begun exploring the use of AI particularly large language models (LLMs) as tools for generating field prompts, offering feedback on reflective journaling, or analyzing student work. Others have taken a more critical stance, positioning AI itself as a topic for student inquiry, for instance, by comparing AI-generated observations with human ones or by discussing data ethics in the context of citizen science. However, there remains a notable lack of systematic, real-world evaluations of AI-supported learning tools in outdoor settings. The logistics of outdoor environments complicate device use and data tracking, and they also intensify concerns around surveillance, consent, and equitable access. A methodologically grounded review could help the field prepare for thoughtful and ethical integration of AI, one that includes clear reporting on data use, informed consent procedures, and analytic transparency.

Taken together, these developments point to four interconnected reasons why a systematic review of methodological trends in nature-based instruction (NBI) from 2021 to 2025 is both timely and necessary:

1. Post-pandemic recalibration: In the wake of COVID-19, research in NBI adapted by shifting study locations, such as using campus micro-sites and neighbourhood walks and adopting more pragmatic designs like quasi-experiments, as well as lighter, more flexible measurement tools. These changes deserve thorough documentation and critical analysis.
2. The climate and sustainability imperative: NBI offers a powerful platform for teaching about climate change and biodiversity. A closer look at how current studies define and measure outcomes like systems thinking, stewardship, and place-based attachment can help guide the standardization of future assessments.
3. Equity and knowledge justice: The rise of approaches that incorporate Traditional Ecological Knowledge (TEK) and foster community collaboration highlights the need for research

practices that are not only methodologically sound but also ethically grounded. A synthesis can help surface effective models and expose recurring challenges in this space.
4. The rise of AI and learning analytics: Early explorations of AI-driven support and data analytics in outdoor learning environments bring both exciting opportunities and serious concerns. A review can clarify where the evidence is beginning to take shape and where greater caution, reflection, or innovation may be needed.

This review is not intended as a meta-analysis focused on effect sizes. Rather, it is a systematic literature review (SLR) that concentrates specifically on the design of studies, the kinds of measurement tools they employ, the settings (both geographic and institutional) in which they take place, and the types of outcomes they aim to assess. The review limits its scope to empirical, English-language studies published between 2021 and 2025 in higher-tier academic journals (those ranked in the top two quartiles, Q1–Q2), with a preference for open-access publications to support transparency and reuse. By combining descriptive data such as the frequency of different research designs, tools, and regional focuses with a thematic analysis of methodological choices, this review seeks to offer a comprehensive overview of how NBI research is currently being conducted, while also pointing to clear and practical directions for future inquiry.

This review offers three key contributions. First, it outlines the most common research designs currently used in nature-based instruction (NBI) and explores the trade-offs that influence those choices such as practical feasibility, ethical concerns, and the need for ecological validity. Second, it presents a comprehensive overview of the measurement tools researchers are using, including pre/post assessments, Likert-scale surveys, concept inventories, structured observation protocols, and performance-based tasks, along with analytic platforms like SPSS, R, AMOS, NVivo, ATLAS.ti, PLS-SEM, and various learning analytics tools. This mapping highlights where greater standardization could improve consistency across studies and allow for more robust meta-analytic synthesis. Third, the review identifies several critical gaps: a limited number of randomized controlled trials and long-term follow-up studies; inconsistent reporting on instructional fidelity; underrepresentation of Global South contexts, despite the wealth of local ecological knowledge; and the early, largely unexplored integration of AI in outdoor learning. Together, these insights lay the groundwork for a future research agenda that is both methodologically rigorous and sensitive to local context. We provide the first systematic mapping of methodological practice in nature-based instruction (2021–2025), applying PRISMA 2020 and MMAT for rigorous screening and appraisal. Our contribution is a usable synthesis of designs, measures, and analytic approaches, plus a roadmap for better reporting, greater equity, and careful integration of AI

The review is grounded in several key theoretical commitments that help shape its contributions. Nature-based instruction (NBI) is rooted in constructivist and socio-constructivist theories, which see learning as an active process of meaning-making, shaped through tools, dialogue, and community practices. It also draws from experiential learning traditions that emphasize a cycle of hands-on experience, reflection, conceptual understanding, and experimentation. Place-based education offers a valuable perspective for understanding how local ecologies, histories, and cultural contexts provide relevance and shape identity. In turn, frameworks such as culturally sustaining pedagogy and funds of knowledge emphasize the legitimacy of community-held knowledge including Traditional Ecological Knowledge (TEK). Finally, the review draws on principles of assessment for learning, focusing on low-stakes, formative, and authentic tasks that assess not just outcomes, but also the learning process itself. These theoretical orientations inform the review's methodological focus which prioritizing authentic inquiry, rich and varied forms of evidence, and multiple ways of knowing, while still aiming for methodological rigor and the ability to synthesize findings across studies.

To guide its scope, this review sets out a series of research questions:
- RQ1 : What research designs and methods of data collection and analysis are most commonly used in empirical studies on NBI from 2021 to 2025?
- RQ2 : What teaching strategies and instructional approaches define nature-based instruction in science education during this period?
- RQ3 : What types of student outcomes are most often studied, and how are they assessed?
- RQ4 : Where are these studies taking place (e.g., school gardens, forests, coastal areas, urban green spaces), and what forms of delivery are used (formal, informal, or non-formal)? How long and how intensive are the interventions?
- RQ5 : Which levels of education are most represented in the research?

Rather than focusing solely on learning outcomes, this review emphasizes the underlying research architectures which are the designs, tools, and contexts that shape how we understand NBI. The goal isn't to promote a one-size-fits-all model, because such a model wouldn't suit the wide range of settings and communities involved in nature-based education. Instead, the review aims to offer a coherent overview of current practices, a shared vocabulary for discussing measurement and reporting, and a thoughtful set of priorities to inform the next wave of research. Research that is not only methodologically sound but also grounded in place, equity, and relevance.

Methodology

This study is structured as a Systematic Literature Review (SLR), following the guidelines outlined in the PRISMA 2020 framework for transparent and rigorous reporting. From the outset, the review was guided by six research questions (RQ1–RQ6), each addressing a different aspect of nature-based instruction (NBI): research designs and methods, pedagogical approaches, outcome measures, learning contexts, educational levels and geographic regions, and methodological gaps. The review protocol laid out clear procedures in advance, including: the sources of information and search terms used; criteria for including or excluding studies; steps for screening and selecting relevant literature; rules for coding and organizing data; methods for evaluating study quality; and the approach to synthesis and sensitivity analysis. All analytical decisions were carefully recorded to ensure the work could be reproduced, audited, and verified by others.

To gather relevant literature, we conducted searches across two well-established, multidisciplinary databases known for their strong coverage of both education and environmental education:
- Scopus (Elsevier): Searches were performed across article titles, abstracts, and keywords (TITLE-ABS-KEY).

- Web of Science Core Collection (Clarivate): Topic searches (TS) included titles, abstracts, author keywords, and Keywords Plus, spanning the SCI-EXPANDED, SSCI, and ESCI indices.

The search covered publications from January 1, 2021, to September 30, 2025, and was limited to documents in English. We focused on peer-reviewed articles and early-access papers. While review articles were included in the initial search to ensure broad coverage, they were excluded from the final dataset if they were not empirical in nature.

Search Strategy and Strings

The search strategy was developed in stages and structured around four key conceptual categories:
1. the context of science education,
2. terminology related to nature-based instruction (NBI),
3. indicators of empirical research, and
4. constraints related to time and language.

To ensure consistency across databases, the field settings were aligned using TITLE-ABS-KEY for Scopus and TS (topic search) for Web of Science.
Here is the final search string used for Scopus:
("science education" OR "STEM education" OR "science teaching")
AND ("nature-based" OR "outdoor learning" OR "outdoor education"
OR "place-based" OR "field-based" OR "school garden"
OR "green schoolyard" OR "forest school" OR "citizen science"
OR "traditional ecological knowledge" OR "TEK")
AND (study OR trial OR experiment OR quasi-experiment OR "mixed-methods"
OR case OR evaluation OR impact OR effectiveness OR efficacy)
AND PUBYEAR > 2020
AND (LIMIT-TO (LANGUAGE, "English"))

And for the Web of Science Core Collection:
TS=("science education" OR "STEM education" OR "science teaching")
AND TS=("nature-based" OR "outdoor learning" OR "outdoor education"
OR "place-based" OR "field-based" OR "school garden"
OR "green schoolyard" OR "forest school" OR "citizen science"
OR "traditional ecological knowledge" OR "TEK")
AND TS=(study OR trial OR experiment OR quasi-experiment OR "mixed-methods"
OR case OR evaluation OR impact OR effectiveness OR efficacy)
AND LA=(English)
AND PY=(2021–2025)

Every search run was carefully documented, this included the date and time of the searches, the number of results returned, and the export settings used. These details are available in the search log as part of the project's data and materials archive.

Record Management and De-duplication
We collected detailed metadata for each study, including information such as the authors, title, abstract, keywords, year of publication, journal name, DOI or URL, document type,

author affiliations, countries or regions (when available), open-access status, and citation counts. To ensure the dataset was clean and non-redundant, we applied a two-step de-duplication process:

1. DOI Matching: We first standardized DOIs by removing URL prefixes and labels, then matched records based on these normalized identifiers.
2. Title and Year Matching: For records without DOIs, we normalized titles and publication years, converting text to lowercase and removing extra spacing to identify duplicates.

When duplicates were found, we kept the version that contained the most complete metadata. A log of all matched pairs and the retained entries is stored as part of the project archive for transparency and future reference.

Eligibility Criteria

Inclusion criteria for the review were clearly defined. Studies had to meet the following conditions:

1. Be empirical in nature whether using quantitative, qualitative, or mixed-methods approaches and explicitly focus on nature-based instruction (NBI) within science education, either in formal school settings or school-linked environments.
2. Provide sufficient methodological detail regarding research design, implementation, measurement tools, and analysis to allow for meaningful extraction and comparison.
3. Be published in English.
4. Appear in journals ranked in the top two quartiles (Q1 or Q2) at the time of publication, as determined by Scimago SJR or Clarivate JCR; if the two sources differed, a Q1–Q2 ranking in either was considered acceptable.
5. Open access was preferred to support accessibility and verification, though not mandatory.

Exclusion criteria included editorials and commentaries without systematic methods, technical papers on tools or instruments that had not been tested in school or classroom contexts, studies unrelated to science education, and any publications released prior to 2021.

Study Selection (Screening)

The screening process was carried out in two stages, with each stage reviewed independently by two researchers:

- Level 1 – Title and Abstract Review: At this stage, studies were filtered based on their relevance to science education and whether they explicitly addressed nature-based instruction (NBI). Additional filters included evidence of empirical research (such as the presence of participants, data, or analysis), publication year (2021–2025), and language (English).
- Level 2 – Full Text Review: In this round, each study was examined in full to ensure it met the inclusion criteria. This included verifying the study’s methodological detail, confirming the journal’s quartile ranking, and assessing whether the study aligned with the focus of the review. Reasons for excluding studies were recorded to ensure transparency.

The initial pool of studies came from SCOPUS and Web of Science exports. These were filtered by publication year, language, and deduplicated using DOI and title matching. To identify relevant empirical studies in NBI, keyword heuristics were applied across titles, abstracts, and keywords.

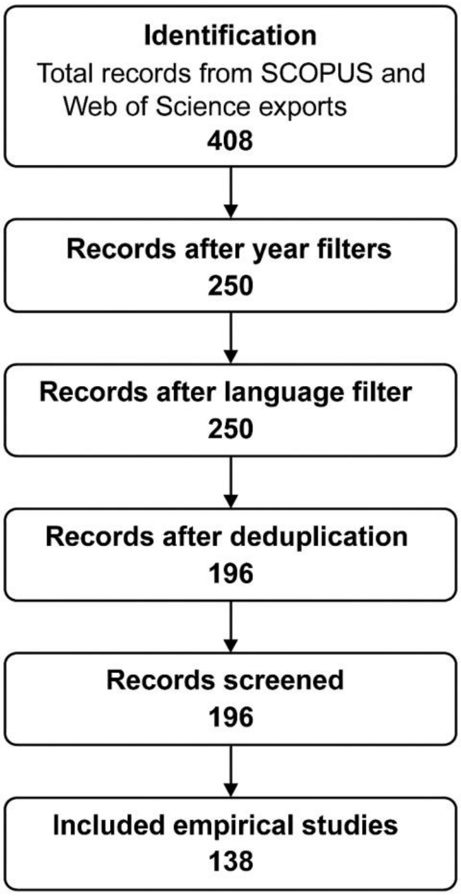

Figure 1. PRISMA Flowchart

Quality Appraisal Using the Mixed Methods Appraisal Tool (MMAT)

Appraisal Approach: To assess the quality of studies included in this review, we applied the Mixed Methods Appraisal Tool (MMAT) to all those that passed the PRISMA screening. Each study was categorized based on its design (qualitative, randomized quantitative, non-randomized quantitative, descriptive quantitative, or mixed methods). Because some records only included abstracts rather than full-text papers, we took a cautious approach: if key evidence was missing, we marked the study as "Can't tell," and if the evidence was limited or ambiguous, we labelled it "Include with reservations." These provisional decisions should be revisited and verified using the full texts before any meta-analytic synthesis.

Summary of MMAT Decisions: Out of the total set of screened studies:

- 33 were fully included,
- 25 were included with reservations,
- 80 were excluded.

More detailed breakdowns by MMAT category and decision type are available in the corresponding table.

Table 1

MMAT Category	Decision	n
Mixed methods (5)	Include	7
Qualitative (1)	Exclude (insufficient reporting)	35
Qualitative (1)	Include	9
Qualitative (1)	Include with reservations	2
Quantitative descriptive (4)	Exclude (insufficient reporting)	40
Quantitative descriptive (4)	Include	16
Quantitative descriptive (4)	Include with reservations	22
Quantitative non-randomized (3)	Exclude (insufficient reporting)	4
Quantitative non-randomized (3)	Include	1
Quantitative randomized (2)	Exclude (insufficient reporting)	1
Quantitative randomized (2)	Include with reservations	1

Methodological Considerations: Studies labelled as “Include with reservations” often had one or more methodological weaknesses. The most common issues were incomplete reporting on how consistently the intervention was delivered (fidelity) or how much exposure participants received (dosage). Others provided limited or no evidence supporting the reliability or validity of their measurement tools. In the case of randomized studies, many lacked clear information about how randomization was carried out or whether groups were comparable at the start. For non-randomized studies, inadequate control for confounding variables was a frequent concern. These limitations highlight the importance of reviewing full texts carefully before including such studies in any quantitative synthesis.

Operational Definitions and Coding Scheme

To guide data extraction and ensure consistency, we developed a detailed codebook in advance, aligning each construct with one of the six research questions (RQ1–RQ6):

- RQ1 – Designs & Methods: We categorized the primary research design used in each study (e.g., randomized experiment, quasi-experiment, mixed-methods, survey, case study, design-based research, ethnography, action research, or observational study). We also documented data collection methods such as interviews, observations, pre/post assessments, Likert scales, concept inventories, performance tasks or artifacts, observation protocols like RTOP/COPUS, and digital tools like logs or sensors. For analytic techniques, we recorded methods such as thematic analysis, statistical packages (e.g., SPSS, R, AMOS), or structural equation modelling tools like PLS-SEM. We required explicit evidence of random assignment for a study to be classified as a randomized trial, and for quasi-experimental designs, a clear pre/post comparison with a control group.
- RQ2 – Pedagogies: We coded for instructional strategies, including outdoor or field-based inquiry, place- and community-based learning, school gardens and campus micro-sites, citizen science, forest schools, experiential models, TEK and culturally sustaining pedagogy, project- or problem-based learning, service learning or stewardship, guided inquiry with protocol, and AI-supported scaffolding.
- RQ3 – Outcomes & Measures: We tracked the types of outcomes studied, conceptual understanding, inquiry skills, engagement and motivation, self-efficacy, values and attitudes, environmental behaviour and stewardship, collaboration and social-emotional learning (SEL), and data or digital literacy. We also noted the instruments used to assess these outcomes, such as tests, inventories, Likert scales, interviews, observations, performance tasks, and analytics or sensor-based data.

- RQ4 – Contexts: We identified the learning environments (e.g., forests, parks, rivers, coastal areas, urban green spaces, community gardens, school campuses, or agricultural settings), the type of delivery (formal, non-formal, or informal education), and the duration or intensity of the intervention (single-day, unit-length, or longitudinal).
- RQ5 – Levels & Regions: We recorded the educational level targeted (primary, lower or upper secondary, tertiary/undergraduate, or teacher education (both preservice and in-service) and geographic region (based on author affiliations or study location, when available).
- RQ6 – Gaps & Trajectories: We noted whether studies provided evidence of causal identification, included longitudinal follow-up, reported on fidelity or use of protocols, used standardized measurement tools, or incorporated AI and analytics.

When labels or classifications were unclear, we referred to detailed examples and justification notes within the codebook to guide consistent decision-making.

Synthesis and Analysis

We used a convergent integrated synthesis approach to align our analysis with the six research questions (RQs). This included:

- Descriptive statistics to summarize the frequency and proportion of different research designs, data collection and analysis methods, pedagogical approaches, measured outcomes, learning settings, and educational levels. These data were visualized using bar charts, line graphs, and radar plots.
- Temporal trend analysis (2021–2025) to track shifts over time, such as increased adoption of mixed-methods designs, emerging methodological combinations, and the growing presence of AI or learning analytics in the research.
- Cross-tabulations to explore how key variables intersect, for example, how pedagogical strategies relate to outcome types, how learning settings correspond to different measurement tools, and how education levels align with study design choices.
- Thematic analysis involved a blend of inductive and deductive coding focused on methodologically relevant excerpts. This helped identify underlying rationales (e.g., the pursuit of ecological validity), recurring constraints (e.g., safety or access issues), and patterns in reporting practices (e.g., fidelity documentation and data governance).

All tables and figures were generated through scripted workflows based on the extracted dataset, with both inputs and outputs systematically versioned to ensure traceability and reproducibility.

Findings

This findings includes only the studies that met the quality threshold under the MMAT screening process, those marked as “Include” or “Include with reservations.” The numbers presented below reflect how these studies were categorized based on the information available in their titles, abstracts, and keywords.

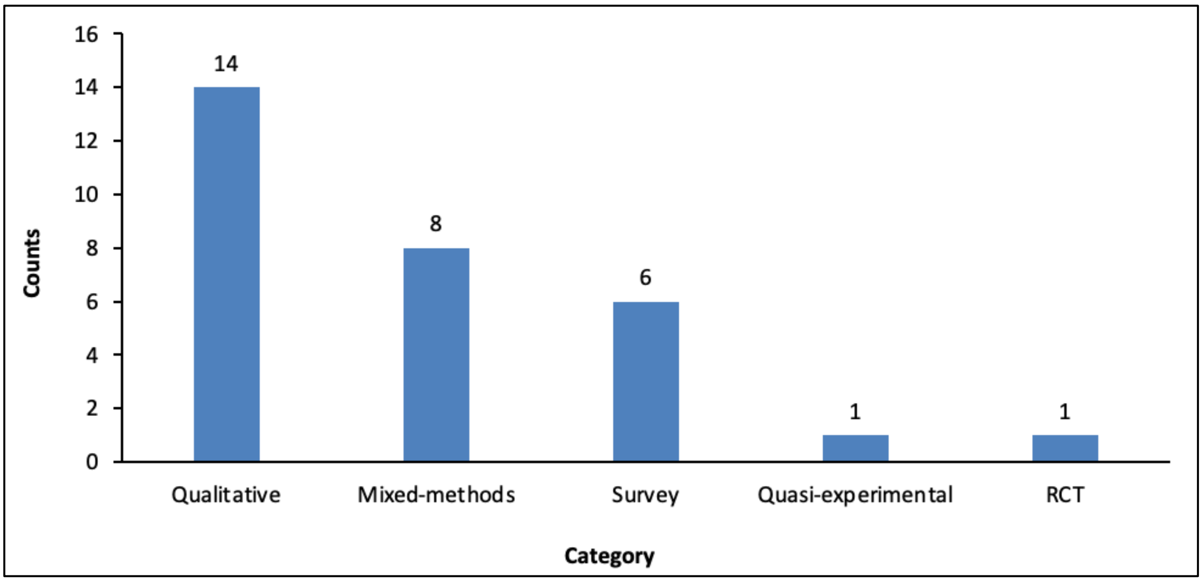

Figure 2. Research Question 1 (RQ1): Research Designs

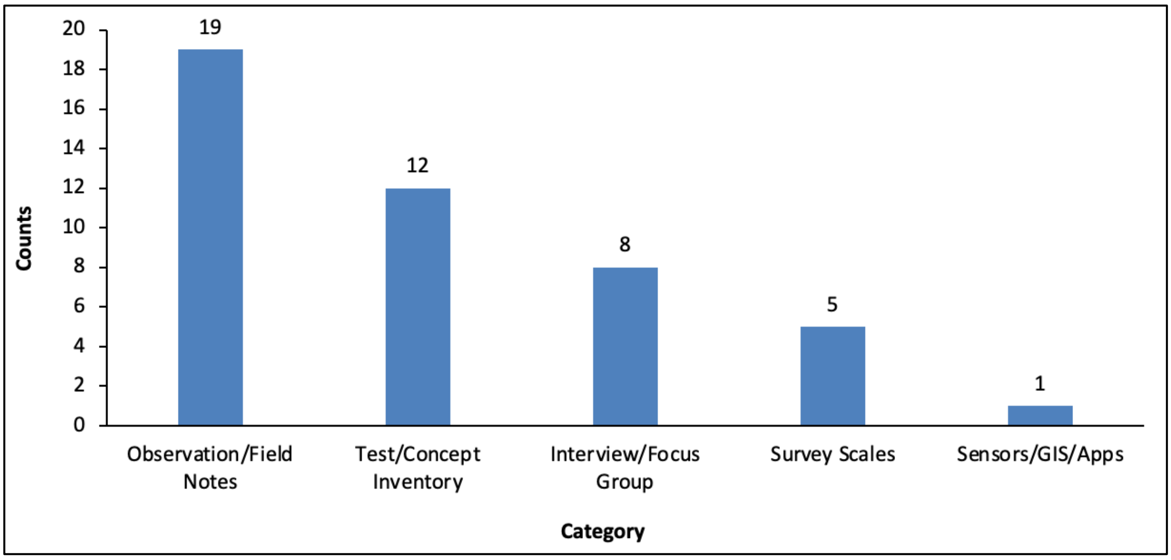

Figure 3. Research Question 1 (RQ1): Data-Collection Techniques

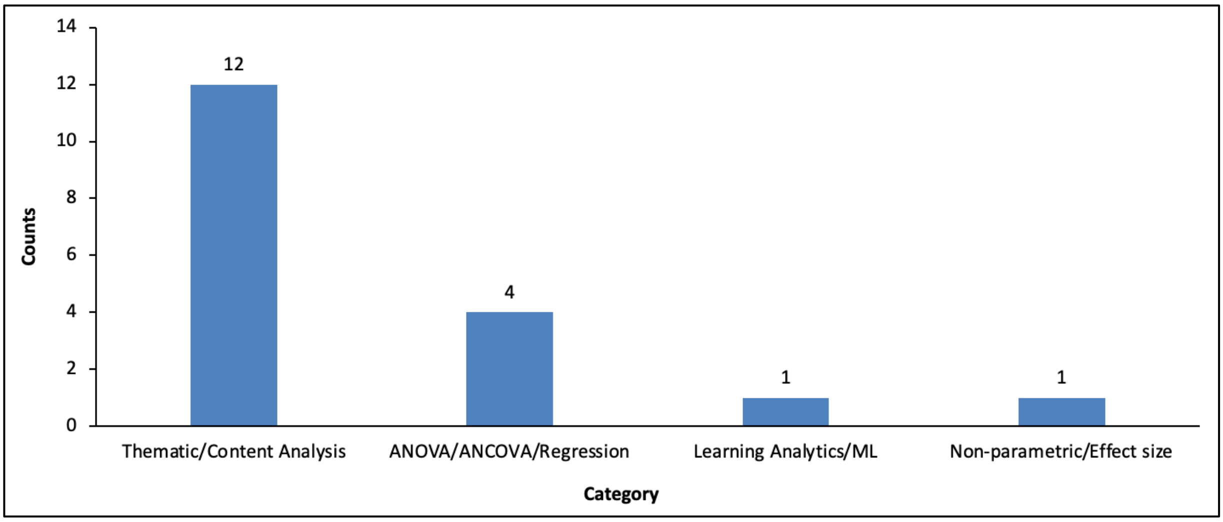

Figure 4. Research Question 1 (RQ1): Analysis Techniques

The studies that passed the MMAT quality screen were predominantly qualitative (n = 14) or used mixed-methods designs (n = 8). A smaller subset employed survey-based approaches (n = 6), and only two studies used designs with stronger causal inference: one quasi-experimental and one randomized controlled trial. This distribution is consistent with recent reviews, which describe the evidence base for nature-based instruction (NBI) as promising but methodologically varied. Well-powered randomized trials that can rigorously isolate treatment effects in real-world school settings remain rare (Mann et al., 2022; Vasilaki & Doulkeridou, 2025).

There are some signs of progress such as the emergence of cluster-randomized trials in early childhood education (e.g., the PRO-ECO project) but these remain limited in number and are often concentrated in specific regions (Ramsden et al., 2025). Within this broader methodological landscape, observations and field notes were the most commonly used forms of data collection (n = 19), followed by tests or concept inventories (n = 12), and interviews or focus groups (n = 8). Survey tools (n = 5) and digital or sensor-based methods (n = 1) were much less frequently used.

In terms of analysis, thematic and content analysis dominated (n = 12), while statistical approaches like ANOVA, ANCOVA, or regression were used in only a few cases (n = 4). More advanced techniques such as machine learning or learning analytics appeared only sporadically (n = 1). These patterns reflect ongoing recommendations in the field to combine rich, descriptive documentation of learning processes with validated outcome measures and analytic strategies that account for clustering and potential confounds (Mann et al., 2022; Vasilaki & Doulkeridou, 2025).

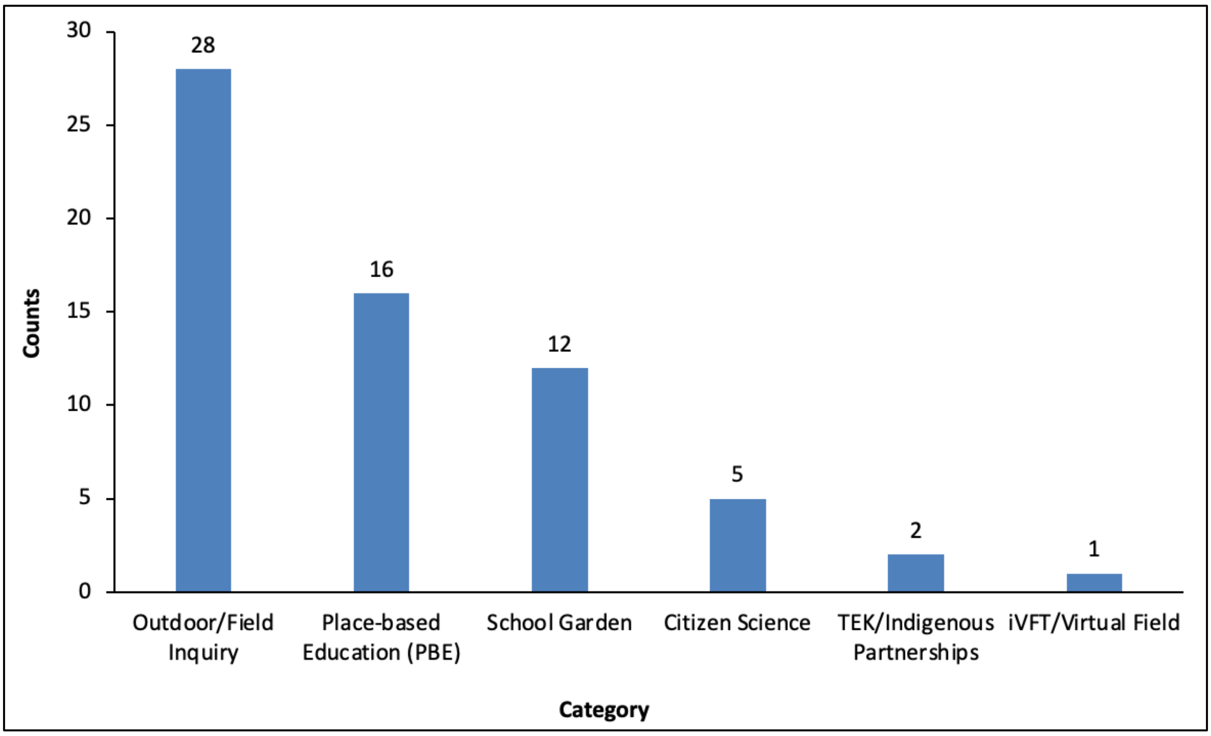

Figure 5. Research Question 2 (RQ2): Pedagogical Approaches

In terms of teaching strategies, the majority of studies focused on outdoor or field-based inquiry (n = 28) and place-based education (PBE) (n = 16). School gardens were also a frequent setting (n = 12), while citizen science projects (n = 5), partnerships involving Traditional Ecological Knowledge (TEK) or Indigenous communities (n = 2), and immersive virtual field trips (iVFT/VR) (n = 1) were less commonly used.

Importantly, the literature emphasizes that simply taking learning outside is not enough to ensure meaningful educational gains. The most effective outcomes occurred when instructional designs intentionally connected real-world phenomena with core scientific practices such as measurement, modelling, and argumentation and paired them with aligned assessment strategies (Mann et al., 2022).

Citizen-science programs were most impactful when students were explicitly taught how to handle data: including quality control, reasoning under uncertainty, and making evidence-based claims (Peltoniemi et al., 2023). Similarly, while immersive technologies like virtual field trips are still emerging, they have shown promise in supporting learning particularly when the experiences are tightly linked to conceptual goals and include structured opportunities for reflection (Vasilaki & Doulkeridou, 2025).

Across all instructional formats, one constant appears: ongoing professional development for teachers, especially in areas like risk management, facilitating inquiry, and formative assessment is a key factor in successful implementation.

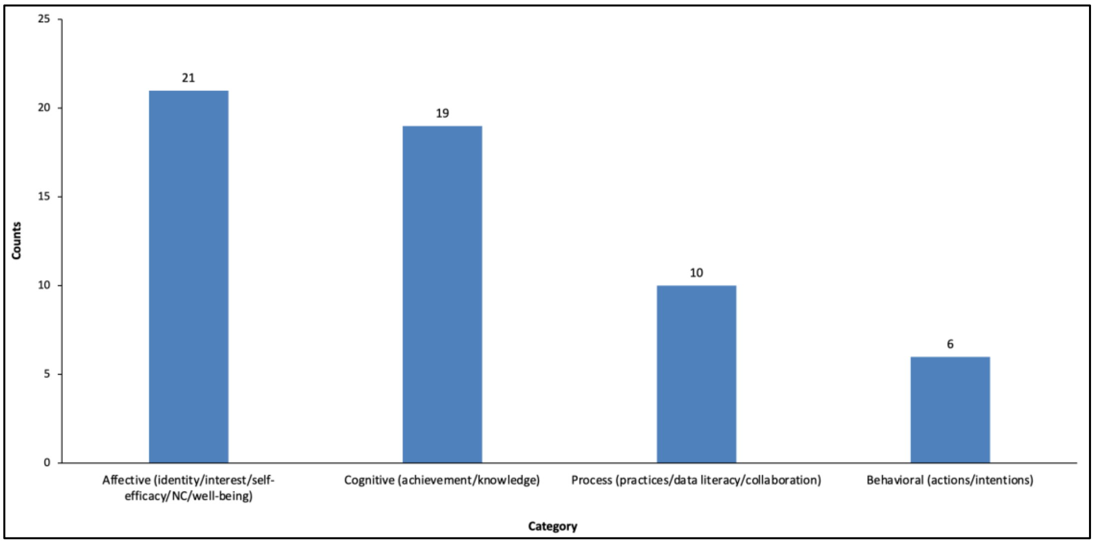

Figure 6. Research Question 3 (RQ3): Targeted Outcomes

When it comes to learning outcomes, most studies focused on affective dimensions such as students' identity, interest, self-efficacy, connection to nature, and overall well-being (n = 21) followed closely by cognitive outcomes, including academic achievement and content knowledge (n = 19). Process-related outcomes like scientific practices, data literacy, and collaboration were less commonly assessed (n = 10), as were behavioural outcomes, such as actions taken or intentions to act (n = 6). However, both of these areas show signs of increasing attention in recent work.

Evidence from meta-analyses and large-scale programs suggests that cognitive gains depend heavily on how well the learning experience is aligned with clear goals and how much exposure students receive. Affective outcomes while consistently positive, require further support through behavioural or performance-based measures to confidently claim long-term impact or transfer (Vasilaki & Doulkeridou, 2025).

Some of the most compelling findings come from multi-site studies involving school gardens. For instance, the TX Sprouts program showed measurable improvements in academic performance (e.g., fourth-grade reading) and student diet when the program was implemented in a sustained and coherent way (Davis et al., 2023). Similarly, a two-year randomized trial in low-income elementary schools found that increased fruit and vegetable consumption occurred only when students received ample exposure and the program was delivered with fidelity (Wells et al., 2022).

These findings underscore the importance of reporting not just outcomes, but also fidelity of implementation and dosage, to help explain variation in results across studies.

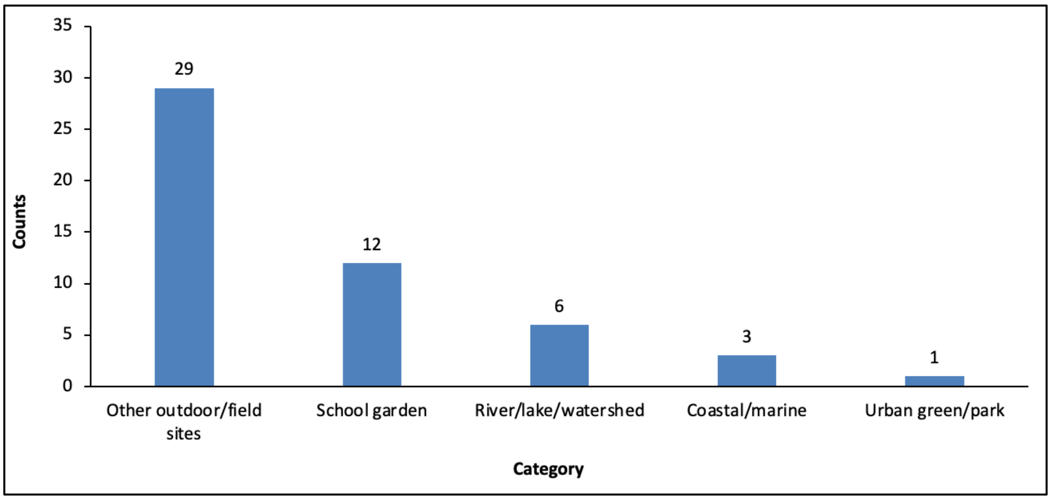

Figure 7. Research Question 4 (RQ4): Settings

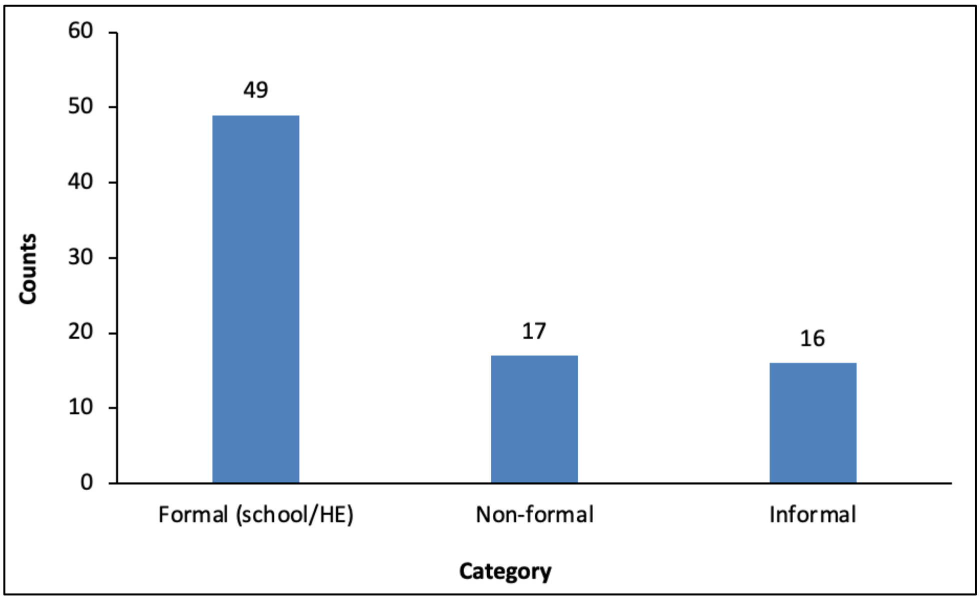

Figure 8. Research Question 4 (RQ4): Delivery Contexts

The settings and delivery formats used in these studies largely reflect the practical realities of implementing nature-based instruction. Most interventions took place in nearby outdoor or field environments (n = 29) and school gardens (n = 12). Fewer studies were set in river or watershed areas (n = 6), coastal or marine environments (n = 3), or urban parks and green spaces (n = 1).

In terms of how learning was delivered, the majority of studies were situated within formal education settings, such as schools or universities (n = 49). However, there was also meaningful engagement through non-formal (n = 17) and informal (n = 16) educational programs.

Consistent with findings from randomized trials, the most lasting cognitive and emotional outcomes were associated with sustained exposure, for example, weekly sessions or semester-long modules that allowed for ongoing measurement, reflection, and modelling. In contrast, one-off activities, even when highly engaging, generally resulted in more modest knowledge gains unless they were part of a broader, coherent learning sequence (Davis et al., 2023; Wells et al., 2022).

Given this, the field would benefit from standardized reporting on how often and for how long students are engaged, the structure of inquiry cycles, and contextual factors such as access to green spaces, safety protocols, and teacher training. Including this kind of detail would significantly improve both the transferability of findings and the ability to assess cost-effectiveness (Mann et al., 2022).

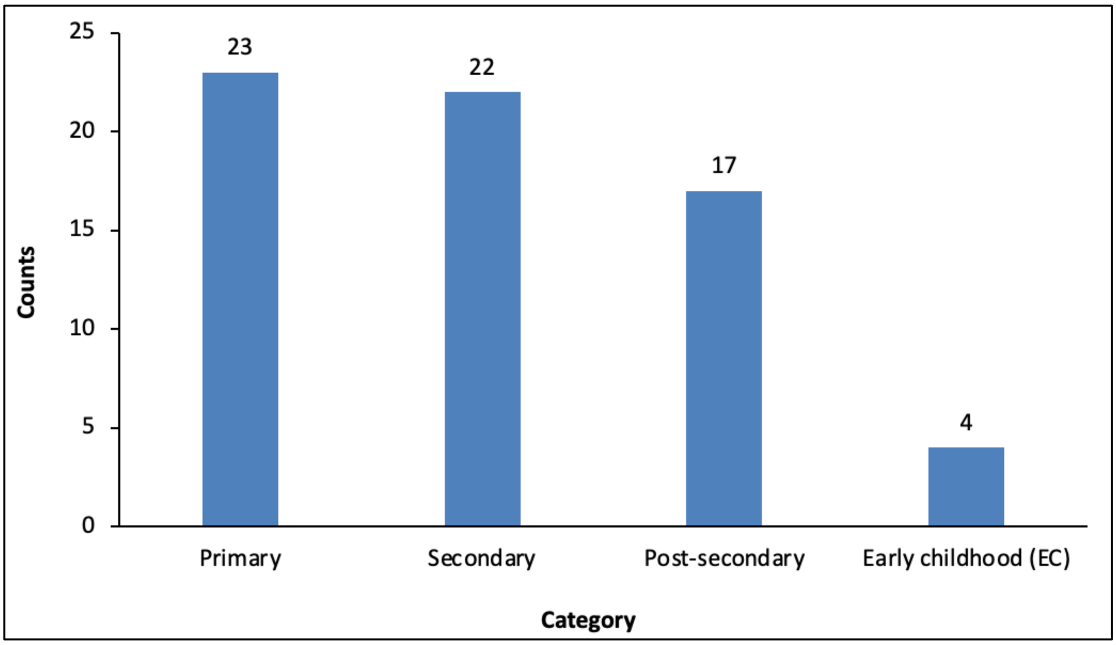

Figure 9. Research Question 5 (RQ5): Educational Levels

Finally, the distribution of studies across different educational levels was fairly balanced. There was strong representation at both the primary level (n = 23) and secondary level (n = 22), along with a significant number of studies in post-secondary or higher education contexts (n = 17). While fewer in number, studies focused on early childhood education (n = 4) are beginning to emerge and show promising growth.

This pattern aligns with recent reviews, which note that most research in nature-based instruction has been concentrated in primary and secondary education, with a new and expanding body of work now focusing on young children particularly in relation to themes like nature connectedness and well-being (Vasilaki & Doulkeridou, 2025).

Discussion

This systematic review analyzed 58 empirical studies on nature-based instruction (NBI) in science education, published between 2021 and 2025. The focus was methodological, looking closely at how researchers designed their studies, what kinds of measures they used, and how they approached data analysis. The goal wasn't to summarize the learning outcomes themselves, but rather to understand broader patterns in how this research is being conducted: What types of study designs are being chosen, under what circumstances, for what reasons, and how well are these decisions being documented in ways that support synthesis, promote equity, and enable thoughtful, responsible scaling?

The outcomes reported in the reviewed studies show a strong emphasis on affective dimensions such as identity, interest, self-efficacy, connection to nature, and wellbeing alongside cognitive outcomes like achievement and conceptual understanding. This balance reflects a more mature theory of change in nature-based instruction: affective engagement and experiences that resonate with students' identities are not just side benefits, but often serve as key drivers of long-term participation in scientific learning.

However, the field still faces challenges in tracking outcomes over time. Many studies rely on short follow-up periods and self-reported behavioural data, which limits our understanding of longer-term impacts. Large, multi-site randomized studies offer a promising path forward. For example, both the TX Sprouts program and a two-year, multi-state randomized trial of school gardens have shown that consistent exposure, program coherence, and faithful implementation are strong predictors of student outcomes and help explain differences between sites (Davis et al., 2023; Wells et al., 2022).

These studies also highlight reporting practices that would benefit the broader field such as clearly documenting dosage (frequency × duration), adherence to core components, and the extent of teacher training. These elements are essential for interpreting mixed results and conducting meaningful cost-effectiveness analyses.

Another important theme that cuts across the literature is the inconsistency in how learning is measured. Many studies rely on teacher-made tests or custom-designed scales that lack evidence of reliability or construct validity. This limits the usefulness of their findings, especially when trying to compare across studies or include them in a meta-analysis. In contrast, research that pairs validated concept inventories and affective scales with performance-based tasks such as rubric-scored data notebooks or student-produced arguments offers both greater depth and better comparability.

Several reviews recommend moving the field toward a "measurement commons": a shared set of openly licensed, age-appropriate tools that assess core science concepts, as well as validated affective measures for constructs like identity, interest, self-efficacy, and nature connectedness. These tools should be supplemented by transparent rubrics that account for reliability and are designed to assess scientific practices and data literacy (Mann et al., 2022; Vasilaki & Doulkeridou, 2025).

Building such a commons would make it easier to aggregate findings across studies, conduct more nuanced analyses (e.g., comparing by setting, grade level, or program duration), and support better-informed planning for large-scale studies, including those using cluster-randomized designs.

The locations and types of institutions represented in the studies mostly formal school settings using nearby outdoor spaces or school garden that closely reflect the logistical and policy constraints that real-world schools face. Encouragingly, some of the strongest research designs are emerging in early childhood education (Ramsden et al., 2025), showing that it is entirely feasible to run rigorous trials when ethical considerations are carefully addressed and randomization is done at the classroom level. Similar investments in underrepresented regions would not only strengthen the generalizability of findings but also promote greater equity in the field.

Taken together, these findings point to several clear priorities for future research.
First, the field needs to deepen its use of causal designs. This means expanding the use of pre-registered cluster-randomized trials and, where randomization isn't possible, employing strong quasi-experimental alternatives such as regression discontinuity or synthetic control methods. These studies should include pre-study power calculations, use appropriate statistical techniques (like multilevel models or cluster-robust estimators) to handle data clustering, and transparently report issues like attrition and missing data (Ramsden et al., 2025).

Second, the community would benefit from collaboratively developing a shared set of measurement tools. This could include age-specific concept inventories, validated scales for affective constructs (like interest, identity, and self-efficacy), and well-documented rubrics for assessing process outcomes. Having common tools would reduce variability across studies, make it easier to combine findings through meta-analysis, and speed up the development of a cohesive theory base (Vasilaki & Doulkeridou, 2025; Mann et al., 2022).

Third, implementation science should become standard practice in this research area. Studies should routinely include tools like fidelity checklists, logs of teacher professional development, and records of student exposure and go beyond description by analyzing how these variables influence outcomes. Findings from garden-based RCTs show that this kind of implementation data isn't just background detail, it actively predicts which programs work and why (Davis et al., 2023; Wells et al., 2022).

Fourth, the field should thoughtfully expand both the scope and delivery methods of nature-based instruction. For example, citizen science programs should clearly define the data literacy skills they aim to teach, design tasks and assessments to match those goals, and document how students engage with the process, including the quality and accuracy of their data workflows (Peltoniemi et al., 2023).

Emerging technologies like immersive virtual field trips (iVFT) and virtual reality (VR) offer exciting possibilities for increasing access and giving students more frequent practice. However, these tools must be evaluated with the same level of rigor as in-person outdoor learning. This includes using standardized outcomes, clear analysis plans, and designs that account for the fact that learning happens in complex, often clustered environments (Vasilaki & Doulkeridou, 2025).

Fifth, the principles of equity and local relevance need to move beyond good intentions and become central to the design of studies. Research partnerships in Africa, Latin America, and other underrepresented regions should be supported with shared tools, phenomena that reflect local contexts, and strong professional development networks for teachers. This will allow researchers to examine how cultural, ecological, and infrastructural differences influence outcomes (Mann et al., 2022).

Finally, as tools like learning analytics and lightweight AI start to appear in outdoor learning such as using image recognition for biodiversity logs or natural language processing (NLP) for analyzing field notes and it's essential to build in ethical safeguards. That means adopting privacy-by-design frameworks, conducting regular bias audits, and documenting how the models work. Most importantly, researchers should evaluate whether these technologies support rather than replace the vital role of teachers and facilitators in guiding scientific practice.

In summary, this review reinforces that nature-based instruction (NBI) is not just about changing the setting, it represents a deeper shift in how science is taught and learned. When instructional approaches are thoughtfully designed to connect real-world phenomena with hands-on, data-rich inquiry, and when assessments are clearly aligned with the skills and understandings being targeted, students show meaningful progress, progress that is increasingly measurable and backed by evidence.

What the field needs now is a coordinated methodological effort: stronger approaches to identifying causal effects, shared tools for measuring learning, greater transparency around how programs are implemented, and a broader, more equitable inclusion of regions and communities that have historically been underrepresented.

Taken together, these steps will help turn promising results into reliable, transferable knowledge about what works, for whom, and under what conditions NBI can make the biggest impact in science education (Mann et al., 2022; Vasilaki & Doulkeridou, 2025; Ramsden et al., 2025; Davis et al., 2023; Wells et al., 2022; Peltoniemi et al., 2023).

Conclusion

This review brings together peer-reviewed research on nature-based instruction (NBI) in science education from 2021 to 2025. The evidence reveals a field that is pedagogically rich but methodologically inconsistent. Most studies rely on qualitative or mixed-methods designs, often grounded in extensive use of observations and field notes. Only a small number of studies primarily quasi-experimental and randomized trials which offer stronger causal insight into NBI's effects. This imbalance does not undermine the value of qualitative work, but it highlights the need for a broader mix of rigorous designs to build a more robust evidence base.

Instructionally, the most effective approaches go far beyond simply holding lessons outdoors. Successful programs deliberately integrate local phenomena with core scientific concepts and authentic practices such as data collection, modelling, analysis, and scientific argumentation. These approaches are most commonly implemented in place-based settings and school garden environments, where real-world relevance is tightly connected to the curriculum.

In terms of learning outcomes, most studies focus on affective domains including identity, interest, self-efficacy, nature connectedness, and wellbeing as well as cognitive outcomes like achievement and conceptual understanding. Attention to process-oriented outcomes, such as data literacy and scientific reasoning, is increasing but still limited. Similarly, few studies explore behavioural indicators of learning. What is clear, is that sustained exposure and fidelity to program design strongly influence student outcomes. Yet, reporting on key factors like program dosage, teacher professional development, and implementation integrity remains inconsistent across the literature.

This review contributes to the field in three meaningful ways. First, it offers a quality-filtered overview of research designs, pedagogical approaches, outcomes, and contexts which helping to clarify where evidence is accumulating and where it is still lacking. Second, it proposes a clear agenda for strengthening measurement and reporting. This includes the use of validated tools for cognitive and affective constructs, well-constructed rubrics for assessing practices like modelling or argumentation, and standardized reporting of implementation variables such as frequency, duration, and teacher support. Third, it identifies the conditions under which NBI is most effective namely, when learners engage in sustained inquiry, participate in data-rich tasks, and receive formative feedback in field or simulation-based environments.

Of course, this review has its limitations. The wide variation in assessment tools used across studies made it difficult to compare outcomes directly or synthesize findings quantitatively. Additionally, there is a noticeable geographic skew: most studies were conducted in Europe or North America, with limited representation from the Global South or Indigenous communities.

To address these gaps, future work should take four important directions. First, the field should expand the use of pre-registered, cluster-randomized trials and high-quality quasi-experimental designs, using multilevel analysis to account for complex learning settings. Second, researchers should work together to create a shared “measurement commons” a suite of open-access, age-specific tools and rubrics that promote consistency and enable credible cross-study comparisons. Third, implementation science should become standard practice: studies need to report fidelity, dosage, and teacher development, and examine how these variables affect learning outcomes. Finally, the field must prioritize equity and contextual diversity by investing in partnerships, replications, and locally grounded work in underrepresented regions.

Together, these steps can turn isolated promising findings into reliable, transferable knowledge about how nature-based instruction works, for whom it is most effective, and under what conditions it produces the greatest benefits for science education.

Acknowledgement
This research is funded by grant number GG-2023-015 by Sumitomo Foundation.

References

Ambrosino, C. M., & Rivera, M. A. J. (2022). A longitudinal analysis of developing marine science identity in a place-based, undergraduate research experience. *International Journal of STEM Education*. https://doi.org/10.1186/s40594-022-00386-4

Ambrosino, C. M., & Rivera, M. A. J. (2023). Relevance of science, conceptualization of scientists, and contextualized "failure" as mediators in the development of student science identity. *CBE—Life Sciences Education*. https://doi.org/10.1187/cbe.22-04-0074

Aragón Núñez, L., & Gómez-Chacón, B. (2022). Promoting enquiry skills in trainee teachers within the context of the university ecological garden. *Education Sciences*. https://doi.org/10.3390/educsci12030214

Aloisio, J. M., Roberts, S. J., Becker-Klein, R., Dunifon, S., Lewis, J. D., Clark, J. A. G., Munshi-South, J., & Tingley, K. (2022). Impacts of a near-peer urban ecology research mentoring program on undergraduate mentors. *Frontiers in Ecology and Evolution*. https://doi.org/10.3389/fevo.2022.803210

Ayotte-Beaudet, J. P., Chastenay, P., Beaudry, M. C., L'Heureux, K., Giamellaro, M., Smith, J., Desjarlais, E., & Paquette, A. (2023). Exploring the impacts of contextualised outdoor science education on learning: The case of primary school students learning about ecosystem relationships. *Journal of Biological Education*. https://doi.org/10.1080/00219266.2021.1909634

Bowers, E. P., Larson, L. R., & Parry, B. J. (2021). Nature as an ecological asset for positive youth development: Empirical evidence from rural communities. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.688574

Carson, S., Rock, J., & Smith, J. K. (2021). Sediments and seashores: A case study of local citizen science contributing to student learning and environmental citizenship. *Frontiers in Education*. https://doi.org/10.3389/feduc.2021.674883

Carrier, S. J., Scharen, D. R., Hayes, M. L., Smith, P. S., Bruce, A., & Craven, L. (2024). Citizen science in elementary classrooms: A tale of two teachers. *Frontiers in Education*. https://doi.org/10.3389/feduc.2024.1470070

Chatmaneerungcharoen, S., Jakra, A., & Nielsen, W. S. (2025). Collaborative lesson study: Enhancing grade 2 students' explanations of scientific phenomena through outdoor STEM learning in the context of Phuket Province. *Multidisciplinary Science Journal*. https://doi.org/10.31893/multiscience.2025559

Chen, S., Sermeno, R., Hodge, K., Murphy, S., Agenbroad, A., Schweitzer, A., Tsao, L., & Roe, A. J. (2024). Young children's self-regulated learning benefited from a metacognition-driven science education intervention for early childhood teachers. *Education Sciences*. https://doi.org/10.3390/educsci14060565

Cincera, J., Kroufek, R., Marková, K., Krepelkova, S., & Šimonová, P. (2021). The GLOBE program: What factors influence students' and teachers' satisfaction with science education. *Research in Science and Technological Education*. https://doi.org/10.1080/02635143.2019.1687441

Corbacho, I., Del Barco, M. A., & Muñoz-Losa, A. (2025). Science and plant interest in outdoor learning: Evaluating prospective teachers' experiences with a botanical inquiry trail. *Science Education International*. https://doi.org/10.33828/sei.v36.i1.6

De-Abreu, S., Vogt, K. A., Schreier, A., Fawcett, P., Marchand, M. E., Vogt, D. J., Tulee, M., George, W., Martin, L., & Matsumoto, K. (2022). Teaching holistic environmental thought: A classroom approach. *Thinking Skills and Creativity*. https://doi.org/10.1016/j.tsc.2022.101141

Eden, A., Anderson, J., & Padolf, A. (2025). Miami, my garden: Creating tomorrow's botanists through a botanical garden–high school partnership. *The American Biology Teacher*. https://doi.org/10.1525/abt.2025.87.4-5.234

Foraster, M., Esnaola, M., López-Vicente, M., Rivas, I., Àlvarez-Pedrerol, M., Persavento, C., Sebastian-Galles, N. Ú., Pujol, J., Dadvand, P., & Sunyer, J. A. (2022). Exposure to road traffic noise and cognitive development in schoolchildren in Barcelona, Spain: A population-based cohort study. *PLOS Medicine*. https://doi.org/10.1371/journal.pmed.1004001

Fu, H., Pan, Z., & Wenderott, J. K. (2025). Engaging with clean power: Integrated, hands-on lessons on photocatalytic hydrogen production for high school students. *Discover Education*. https://doi.org/10.1007/s44217-025-00490-x

Gallay, E. E., Flanagan, C. A., & Parker, B. (2021). Place-based environmental civic science: Urban students using STEM for public good. *Frontiers in Education*. https://doi.org/10.3389/feduc.2021.693455

Harris, E. M., & Ballard, H. L. (2021). Examining student environmental science agency across school science contexts. *Journal of Research in Science Teaching*. https://doi.org/10.1002/tea.21685

Höper, J. (2025). Mind the gap between chemistry and outdoor education: Primary and lower secondary preservice teachers' beliefs during an outdoor-based introduction to chemistry in science education. *Journal of Science Teacher Education*. https://doi.org/10.1080/1046560X.2024.2406661

Jones, D. A. (2025). Integrating STEM education in sustained deep rural schools: Innovative strategies for multi-grade, multi-subject classrooms. *Frontiers in Education*. https://doi.org/10.3389/feduc.2025.1569489

Kesler, M., Kaasinen, A., & Kervinen, A. (2024). Teaching science outdoors: Supporting pre-service teachers' skill development with the help of available mobile applications. *Education Sciences*. https://doi.org/10.3390/educsci14111218

Kostøl, K. B., & Remmen, K. B. (2022). A qualitative study of teachers' and students' experiences with a context-based curriculum unit designed in collaboration with STEM professionals and science educators. *Disciplinary and Interdisciplinary Science Education Research*. https://doi.org/10.1186/s43031-022-00066-x

Kutlu-Abu, N., & Dal, E. (2025). Gifted students' perceptions on science courses in general education classroom. *Journal of Baltic Science Education*. https://doi.org/10.33225/jbse/25.24.296

Lee, M., Zhang, R., Shojaei, A., & Roofigari-Esfahan, N. (2025). A human–building interaction (HBI) system for multidimensional occupant feedback integration and predictive modeling. *Journal of Building Engineering*. https://doi.org/10.1016/j.jobe.2025.113839

Ma, H., & Green, M. M. (2021). The affordance of place in developing place-responsive science teaching pedagogy: Reflections from pre-service teachers. *Journal of Science Teacher Education*. https://doi.org/10.1080/1046560X.2021.1898139

Mercier, A. K., Sivitskis, A. J., Torbert, J., & Vallier, K. (2025). Shaping change locally: A place-based STEM project's influence on rural elementary and middle grade students. *Frontiers in Education*. https://doi.org/10.3389/feduc.2025.1584805

Mills, R., Scholes, L. J., Stahl, G. D., McDonald, S., Comber, B. M., & Brownlee, J. L. (2021). How do rural Australian students' ethnogeographies related to people and place influence their STEM career aspirations? *International Journal of Science Education*. https://doi.org/10.1080/09500693.2021.1961326

Molnár, Z., Radács, M., & Gálfi, M. (2024). Fieldwork integration into the primary school curriculum to develop complex scientific thinking. *International Electronic Journal of Elementary Education*. https://doi.org/10.26822/iejee.2024.370

Ogodo, J. A., Nesmith, S. M., Baker, C., Boddie, S. C., Spence, A. M., Nesmith, D., Crumpton, T., Small, A., Brown, A., & Lilley, L. (2023). Think like a scientist and engineer: Schools, university, and community partnerships using an aquaponics project during COVID-19. *Journal of Higher Education Theory and Practice*. https://doi.org/10.33423/jhetp.v23i8.6059

Ortiz-Guerrero, C., & Loizzo, J. L. (2024). Rocks really rock: Electronic field trips via Google Earth can generate positive impacts in attitudes toward Earth sciences in middle- and high-school students. *Geoscience Communication*. https://doi.org/10.5194/gc-7-101-2024

Othman, O., Iksan, Z. H., & Ruhizan, M. Y. M. (2022). Creative Teaching STEM Module: High school students' perception. *European Journal of Educational Research*. https://doi.org/10.12973/eu-jer.11.4.2127

Özyurt, Z., Türkoğlu, I., & Bahçeci, F. (2025). Investigation of biodiversity awareness and conservation behaviors on science teacher candidates. *Challenges in Sustainability*. https://doi.org/10.56578/cis130107

Photo, P. (2024). Exploring the impact of zoological garden visits on science learning in primary school: Acknowledging and transforming prior knowledge via outdoor learning. *Journal of Outdoor and Environmental Education*. https://doi.org/10.1007/s42322-024-00175-3

Pitiporntapin, S., Butkatanyoo, O., Piyapimonsit, C., Thanarachataphoom, T., Chotitham, S., & Lalitpasan, U. (2024). The influence of outdoor STEM activities on students' STEM literacy: A study of elective subjects, extracurricular activities, and student clubs. *Jurnal Pendidikan IPA Indonesia*. https://doi.org/10.15294/es9jaz32

Pitiporntapin, S., Butkatunyoo, O., Piyapimonsit, C., Thanarachataphoom, T., Chotitham, S., & Lalitpasan, U. (2023). The development of a professional development model focusing on outdoor learning resources to enhance in-service teachers' STEM literacy. *Kasetsart Journal of Social Sciences*. https://doi.org/10.34044/j.kjss.2023.44.2.19

Pitiporntapin, S., Butkatunyoo, O., Piyapimonsit, C., Thanarachataphoom, T., Chotitham, S., & Lalitpasan, U. (2025). Understanding teacher practices in outdoor STEM implementation: A pedagogical content knowledge perspective. *Research in Science Education*. https://doi.org/10.1007/s11165-025-10272-y

Puchongprawet, J., & Pitiporntapin, S. (2025). Integrating outdoor STEM activities into an elective course to promote students' STEM literacy. *International Journal of Innovation in Science and Mathematics Education*. https://doi.org/10.30722/33.01.004

Ramsurrun, H., Elaheebocus, R., & Chiniah, A. (2024). Digital tools in informal science education sites: A systematic literature review. *Journal of Science Education and Technology*. https://doi.org/10.1007/s10956-024-10105-z

Rico, A., Agirre-Basurko, E., Ruiz-González, A., Palacios-Agúndez, I., & Zuazagoitia, D. (2021). Integrating mathematics and science teaching in the context of education for sustainable development: Design and pilot implementation of a teaching–learning sequence about air quality with pre-service primary teachers. *Sustainability*. https://doi.org/10.3390/su13084500

Rillero, P., Jiménez-Silva, M., Short-Meyerson, K. J., & Rillero, K. M. (2025). From seeds to harvest in seven weeks: Project-based learning with Latina girls and their parents. *Education Sciences*. https://doi.org/10.3390/educsci15020246

Rollwagen-Bollens, G. C., Holmlund, T. D., & Wait, J. (2022). Actively participating in university-sponsored ecological research increases high school students' knowledge of and attitudes about science. *Frontiers in Environmental Science*. https://doi.org/10.3389/fenvs.2021.797769

Stern, M. J., Powell, R. B., Frensley, B. T., Anderson, K. C., Kiewra, L. A., Pownall, M., Schmitt, T., & Thorpe, E. G. (2025). Principles for better environmental education field trips for early adolescent youth in the United States: An empirical study. *Environmental Education Research*. https://doi.org/10.1080/13504622.2025.2502575

Susman, K., Vošnjak, M., & Pavlin, J. (2024). Evaluating the development of pre-service primary school teachers' competences in the context of a student-centred science visits course. *Education Sciences*. https://doi.org/10.3390/educsci14111219

Toolin, R., & Blouin, M. (2025). Place as context, place as inquiry: A phenomenographic study of teachers' conceptions of place-based education in the CZ-CREST program. *Journal of Geoscience Education*. https://doi.org/10.1080/10899995.2025.2530763

Treibergs, K. A., Esparza, D., Yamazaki, J. A., & Smith, M. K. (2023). Journal reflections shed light on challenges students face in an introductory field biology course. *Ecosphere*. https://doi.org/10.1002/ecs2.4509

Tserej, O. N., Sidoti, B. J., Assael, S., Padolf, A., Bistrain, M., Shen, J., & Feeley, K. J. (2024). How do teachers learn and engage with climate change? An examination of the Shade Our Schools—Leaves Are Cool! citizen science program. *Citizen Science: Theory and Practice*. https://doi.org/10.5334/cstp.619

Uskola, A., & Seijas, N. (2023). Use of data obtained in the field and its contribution to the process of construction of the geological change model by preservice elementary teachers. *Research in Science and Technological Education*. https://doi.org/10.1080/02635143.2021.2005561

Väljataga, T., & Mettis, K. (2022). Secondary education students' knowledge gain and scaffolding needs in mobile outdoor learning settings. *Sustainability*. https://doi.org/10.3390/su14127031

Wieselmann, J. R., Sager, M. T., Scott, C. C., & Brown, S. M. (2025). STEM, sports, and service-learning: Exploring undergraduates' experiential learning and social consciousness as summer program volunteers. *Journal of Experiential Education*. https://doi.org/10.1177/10538259251331584

Wright, G. W., Zangori, L. A., Otto, S., Snyder, R., & Cole, L. B. (2025). Exploring middle school science teachers' curriculum sensemaking. *Journal of Science Teacher Education*. https://doi.org/10.1080/1046560X.2025.2497188

Zucker, T. A., Mesa, M. P., DeMaster, D. M., Oh, Y., Assel, M. A., McCallum, C., & Bambha, V. P. (2024). Evaluation of a community-based, hybrid STEM family engagement program at pre-kindergarten entry. *Frontiers in Education*. https://doi.org/10.3389/feduc.2024.1281161