Vol. 4 No. 2, November, 2026. Page. 69-77
ISSN (e): 3046-9570,
DOI: https://doi.org/10.57176/9nc7wj38
Mail: journal.primer@unu-jogja.ac.id

Development and Validation of a Contextualized Science Learning Module for Grade 4 Learners: *Effects on Conceptual Understanding and Motivation*

Jenena C. Baculio [1]
[1]Graduate School of Education, University of Perpetual Help System DALTA, Las Piñas City, Philippines

Article Info

Article history:

Received September 21, 2026.
Revised September 30, 2026.
Accepted November 02, 2026.

Keywords:

ADDIE model
Conceptual understanding
Elementary science
Learning module
Science motivation
SOLO taxonomy

Copyright ©
Universitas Nahdlatul Ulama Yogyakarta.
All rights reserved.

Corresponding Author:
Jenena C. Baculio
jenena.baculio001@deped.gov.ph

ABSTRACT

This exploratory sequential mixed-methods study developed, validated, and evaluated a contextualized Grade 4 Science learning module intended to strengthen conceptual understanding and motivation. Participants included 50 Grade 4 learners, 30 teachers, and 20 science experts from a public elementary-school context in Iligan City, Philippines. Needs-assessment evidence informed module content and design. The module followed the Analysis, Design, Development, Implementation, and Evaluation model and integrated the Structure of Observed Learning Outcomes taxonomy, inquiry-oriented tasks, familiar local examples, collaborative work, and motivational supports. Expert and teacher review examined content validity, instructional quality, and language appropriateness. A one-group pretest–posttest design assessed learner outcomes. Conceptual understanding improved significantly, with a mean increase of 0.52 Structure of Observed Learning Outcomes level units, $t(49) = -5.987$, $p < .001$, $d = 0.85$. Motivation increased from $M = 1.78$ before implementation to $M = 3.23$ afterward, $t(49) = 12.83$, $p < .001$, $d = 1.81$. Learners and teachers associated improvement with hands-on inquiry, scaffolding, and real-life relevance, while recommending simpler vocabulary, more visual support, and low-cost activity alternatives. The module shows promise as a supplementary science-learning resource in resource-constrained settings. However, the one-group design, single-school setting, and convenience sampling preclude strong causal and population-level conclusions.

Abstrak

Penelitian metode campuran sekuensial eksploratif ini mengembangkan, memvalidasi, dan mengevaluasi modul pembelajaran Sains kelas IV yang dikontekstualisasikan untuk meningkatkan pemahaman konseptual dan motivasi. Partisipan terdiri atas 50 siswa kelas IV, 30 guru, dan 20 pakar sains dalam konteks sekolah dasar negeri di Kota Iligan, Filipina. Bukti analisis kebutuhan digunakan untuk menentukan isi dan desain modul. Modul mengikuti model Analysis, Design, Development, Implementation, and Evaluation serta mengintegrasikan taksonomi Structure of Observed Learning Outcomes, tugas berbasis inkuiri, contoh lokal, kerja kolaboratif, dan dukungan motivasi. Tinjauan pakar dan guru menilai validitas isi, kualitas pembelajaran, serta kesesuaian bahasa. Desain satu kelompok prates-pascates digunakan untuk menilai hasil belajar. Pemahaman konseptual meningkat secara signifikan sebesar 0,52 unit tingkat Structure of Observed Learning Outcomes, $t(49) = -5,987$, $p < 0,001$, $d = 0,85$. Motivasi meningkat dari $M = 1,78$ menjadi $M = 3,23$, $t(49) = 12,83$, $p < 0,001$, $d = 1,81$. Siswa dan guru mengaitkan perbaikan dengan inkuiri langsung, scaffolding, dan relevansi kehidupan nyata. Namun, desain satu kelompok, satu sekolah, dan convenience sampling membatasi kesimpulan kausal dan generalisasi.

1. INTRODUCTION

Developing conceptual understanding in elementary science requires more than the memorization of facts. Grade 4 learners must connect observable events with scientific explanations, integrate related ideas, and apply those ideas to unfamiliar situations. These demands become especially difficult when instruction relies heavily on teacher explanation and when schools have limited access to laboratory and digital resources. International assessments have repeatedly shown that many learners struggle to transfer science knowledge to authentic

ISSN (e): 3046-9570

problems, underscoring the need for instruction that makes reasoning visible and connects content to everyday experience (OECD, 2023).

Motivation is equally important. Learners are more likely to persist when tasks support competence, meaningful choice, and positive relationships (Ryan & Deci, 2020). In science classes, hands-on inquiry, attainable challenges, feedback, and locally familiar examples can strengthen both engagement and understanding. Instructional materials should therefore integrate cognitive progression with motivational support rather than treat achievement and interest as separate concerns.

The present study addressed these needs by developing a contextualized Grade 4 Science module through the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) model (Branch, 2009). The module used the Structure of Observed Learning Outcomes (SOLO) taxonomy to organize tasks from single-aspect recognition toward integrated and extended reasoning (Biggs & Collis, 1982). It also incorporated inquiry, reflection, collaborative work, and real-life applications involving familiar materials and local environmental concerns.

The study makes three contributions. First, it connects needs analysis directly to module content and instructional features. Second, it evaluates both conceptual and motivational outcomes using pre-intervention and post-intervention measures. Third, it combines statistical results with learner and teacher feedback to identify the module's strengths and areas for refinement. This integrated evidence is particularly relevant to public-school settings in which resource limitations require practical and adaptable instructional materials.

Elementary science learning is expected to develop more than factual recall. Learners must recognize patterns, distinguish evidence from opinion, connect causes and consequences, and use scientific ideas when interpreting events in daily life. These expectations are challenging during the middle elementary grades because learners are beginning to move from explanations grounded mainly in immediately observable features toward more coordinated reasoning. When classroom activities emphasize terminology and answer production without asking learners to explain relationships, apparent success may conceal fragmented knowledge that is difficult to transfer.

Large-scale science assessments operationalize scientific literacy as the ability to explain phenomena, evaluate inquiry, and interpret evidence in context (OECD, 2023). This orientation is relevant to Philippine elementary education, where curriculum expectations include inquiry, application, environmental responsibility, and the progressive integration of concepts. Yet schools differ considerably in access to laboratories, devices, printed resources, and teacher preparation time. A practical intervention must therefore preserve conceptual rigor while remaining feasible with familiar, inexpensive materials.

The source dissertation identified recurring difficulty with ecosystems, energy, forces, changes in matter, chemical properties, and biodegradability. These topics require learners to coordinate several ideas and to reason about processes that are not always directly visible. Learners also reported difficulty with technical vocabulary and multi-step tasks. These challenges supported the development of a module that used concrete examples, explicit scaffolds, reflection, and locally recognizable environmental applications.

Prior research has established that active learning, inquiry, formative feedback, and motivational support can improve science learning. Nevertheless, evidence about an approach in general does not automatically establish the quality of a specific module for a specific learner population. Instructional materials must be grounded in diagnosed needs, aligned with curriculum expectations, reviewed for scientific and linguistic appropriateness, and evaluated under the conditions in which they are intended to be used.

The present study addressed a design-and-evaluation gap by connecting three elements that are often treated separately: a needs analysis, a transparent instructional-development process, and the simultaneous assessment of conceptual understanding and motivation. The module was not evaluated solely through satisfaction ratings. Learner performance was classified using a progression from fragmented to integrated understanding, while motivation was examined before and after implementation. Qualitative feedback was then used to explain how learners and teachers experienced the intervention and which revisions remained necessary.

This contribution is deliberately context-specific. It does not claim that one module resolves structural problems in science education or that pretest–posttest gains prove causality. Instead, it evaluates whether a systematically developed, low-resource module was sufficiently valid, usable, and educationally promising to justify refinement and a stronger comparative trial.

This study developed and validated a contextualized Grade 4 Science module and examined its association with learners' conceptual understanding and motivation. It investigated learner and teacher needs, expert and teacher judgments of module quality, pretest–posttest differences in conceptual understanding and motivation, and participant accounts of effectiveness, clarity, usability, and needed improvements. The null hypotheses stated that conceptual understanding and motivation would not differ significantly before and after module use at the .05

ISSN (e): 3046-9570

level. The null hypotheses stated that there would be no significant pretest-posttest difference in conceptual understanding and no significant pre-intervention-post-intervention difference in motivation at the .05 level.

Conceptual understanding involves recognizing the organizing relationships that make facts meaningful and usable. In science, a learner who can name materials or organisms may still be unable to explain why they behave in particular ways, how evidence supports a conclusion, or how a principle applies under changed conditions. Conceptual learning therefore requires the reorganization of prior ideas, not simply the accumulation of correct statements. Conceptual-change research emphasizes that learners interpret new information through existing mental models and may retain misconceptions unless instruction makes competing explanations explicit and intelligible (Posner et al., 1982).

Instructional design for conceptual change should create opportunities to articulate prior ideas, observe or analyze discrepant evidence, compare explanations, and apply a revised model. These opportunities are especially important in elementary science, where everyday language can conflict with disciplinary meanings. Terms such as force, energy, change, material, and living may appear familiar while carrying more precise scientific relationships. Scaffolds must therefore support vocabulary and reasoning together rather than assuming that a definition alone produces understanding.

The National Research Council (2012) similarly describes science learning as the integration of disciplinary ideas with scientific practices. From this perspective, explaining, modeling, investigating, analyzing data, and arguing from evidence are not optional additions to content; they are ways through which content becomes usable knowledge. The module in this study translated that principle into short, guided activities that asked learners to observe familiar materials, compare cases, explain patterns, and connect conclusions to local environmental practices.

The SOLO taxonomy describes increasing complexity in learner responses: prestructural, unistructural, multistructural, relational, and extended abstract (Biggs & Collis, 1982). The module used this sequence to scaffold activities from identifying a relevant idea to linking multiple ideas and transferring an explanation to a novel context. The framework provided both a design logic and an outcome classification for conceptual growth.

The SOLO taxonomy offers a practical lens for describing the structural complexity of learner responses (Biggs & Collis, 1982). At the prestructural level, a response is irrelevant or shows no coherent understanding. A unistructural response uses one relevant aspect, while a multistructural response identifies several relevant aspects without integrating them. Relational responses connect ideas into a coherent explanation. Extended-abstract responses generalize, hypothesize, or transfer the organizing principle to a new context. This hierarchy is useful for module design because it aligns instructional scaffolding with the quality of reasoning expected from learners. SOLO does not imply that every learner moves through a rigid sequence at the same rate or that a single score fully represents understanding. Performance varies by topic, task wording, prior experience, and support. In this study, the taxonomy was used as an instructional and descriptive framework: activities progressed from recognition and description to explanation and application, and responses were summarized by their observed structural level. The interpretation focused on group movement in response complexity rather than treating a taxonomy label as a fixed learner trait.

Self-determination theory explains that motivation is strengthened when learners experience autonomy, competence, and relatedness (Ryan & Deci, 2020). The module supported competence through guided examples and progressive difficulty, autonomy through choices and reflection, and relatedness through paired and group tasks. Motivation was examined through interest, effort and persistence, confidence, and perceived value of science learning.

Motivation influences whether learners initiate tasks, persist through confusion, seek help, and invest effort in revising an explanation. Self-determination theory proposes that high-quality motivation is supported when learning environments foster autonomy, competence, and relatedness (Ryan & Deci, 2020). Autonomy support does not require unstructured choice; it involves meaningful participation, understandable reasons for tasks, and opportunities to exercise agency. Competence grows through optimally challenging work, clear guidance, informative feedback, and visible progress. Relatedness develops when learners experience respect, belonging, and productive collaboration.

Science motivation also includes the perceived value of scientific activity and confidence in completing science tasks. Glynn et al. (2011) showed that science motivation can be examined through dimensions such as intrinsic motivation, self-efficacy, self-determination, and relevance. Instruments designed for older learners cannot simply be transferred to Grade 4 without modification. Accordingly, the present study simplified wording, used a four-point response scale, sought expert review, and pilot-tested the questionnaire with learners who resembled the main participants.

A learning module can support motivation when its activities help learners succeed at meaningful intellectual work. Hands-on activity alone is insufficient if learners manipulate materials without connecting observations to concepts. Conversely, excessive explanation may reduce participation and reinforce dependence on the teacher.

ISSN (e): 3046-9570

The developed module attempted to balance structure and agency through guided experiments, collaborative questions, reflection prompts, gradual difficulty, and familiar applications.

ADDIE supplied the development sequence. Analysis identified misconceptions, priority competencies, and classroom constraints. Design translated those findings into objectives, lesson sequences, assessments, and motivational features. Development produced and validated the module. Implementation involved classroom use, and evaluation combined learning outcomes with stakeholder feedback. This process linked the intervention to documented learner needs rather than beginning with a predetermined set of activities.

The ADDIE model organizes development through analysis, design, development, implementation, and evaluation (Branch, 2009). Its value lies less in a fixed sequence than in maintaining alignment among needs, objectives, activities, assessments, and revision evidence. Analysis in the present study identified difficult concepts and motivational barriers. Design translated those findings into intended outcomes and a progression of tasks. Development produced lesson materials, worksheets, assessments, and supports. Implementation examined classroom feasibility, and evaluation integrated expert judgment, learner outcomes, and stakeholder feedback.

Contextualization connected target concepts with materials and situations familiar to learners in Iligan City, including banana leaves, food waste, paper, metals, plastics, composting, and waste management. Context was used as a bridge to disciplinary reasoning, not as a substitute for it. Activities asked learners to classify, predict, explain, and propose actions based on scientific properties. Low-cost alternatives were emphasized so that module use would not depend on continuous internet access or specialized equipment.

The literature therefore supports an integrated design proposition: conceptual progression should be explicit, motivational needs should be supported through classroom structures, and learning tasks should connect scientific reasoning to recognizable situations. The study tested the feasibility and educational promise of that proposition in one Grade 4 setting.

2. METHODS

Research Design

An exploratory sequential mixed-methods design was used. Qualitative and quantitative needs-assessment evidence first identified conceptual difficulties, motivational barriers, and preferred instructional features. These findings informed module development. A quantitative one-group pretest-posttest phase then evaluated changes in conceptual understanding and motivation, while post-implementation feedback and interviews explained how participants experienced the module. Because the design did not include a control group or random assignment, observed gains cannot be attributed exclusively to the intervention.

Participants and Setting

Participation was phase-specific. Fifty Grade 4 learners from one public elementary school in Iligan City completed the needs assessment, main module implementation, pretest, posttest, motivation measures, and learner feedback. Thirty teachers contributed to the needs assessment, module appraisal, and post-implementation feedback, while 20 science experts evaluated content validity, instructional quality, and language appropriateness. A separate group of 30 nonparticipant Grade 4 learners completed pilot testing and was not included in the main outcome analyses. Thus, the intervention results and paired tests refer only to the 50 learners who completed both measurement occasions. Convenience sampling, use of one school, and the distinct functions of each respondent group limit generalizability

Module Development

Needs-assessment findings indicated difficulties with ecosystems, energy, forces, states of matter, chemical properties, and biodegradability, as well as problems with technical vocabulary and applying concepts to everyday situations. In response, the module incorporated explicit learning objectives, scaffolded SOLO-based activities, inquiry and problem-solving tasks, guided reflection, and examples using locally available materials such as banana leaves, food waste, paper, metal scraps, and common plastics. Pilot feedback led to simplified directions, additional examples and visual supports, shorter tasks, and resource-friendly alternatives.

The analysis phase combined learner questionnaires, teacher input, classroom observations, and diagnostic evidence. Priority was given to competencies that respondents considered important but difficult and to topics for which learners showed fragmented explanations. During design, objectives and assessments were aligned, tasks were sequenced using SOLO levels, and PISA-oriented real-life problems were adapted to the developmental level of Grade 4 learners. During development, the researcher produced lesson guides, learner worksheets, formative checks, reflective prompts, and pretest–posttest measures.

Pilot implementation focused on clarity, timing, material availability, learner participation, and the cognitive demand of tasks. Feedback indicated that some instructions were too complex, several activities required more time than available, and learners needed additional support when moving from listing ideas to explaining relationships. Revisions included shorter directions, guided examples, vocabulary assistance, more visual cues,

ISSN (e): 3046-9570

adjusted pacing, and alternatives using inexpensive materials. The revised module was then used in the main implementation phase

Research Instruments

The study used (1) a needs-assessment questionnaire, (2) a conceptual-understanding test organized by SOLO levels, (3) a four-point science-learning motivation questionnaire, (4) a module-design survey, (5) an expert-validation form, (6) observation and feedback forms, and (7) semi-structured interview guides. Science-education and assessment specialists reviewed the instruments for relevance, clarity, curricular alignment, and age appropriateness. Pilot testing was conducted with nonparticipants. The manuscript records a Cronbach's alpha of .85 for the principal questionnaire used in the main evaluation, indicating satisfactory internal consistency.

The needs-assessment questionnaire used four-point ratings and open comments to identify difficult concepts and desired instructional features. The conceptual-understanding assessment contained selected-response and constructed-response tasks organized to elicit unistructural, multistructural, relational, and extended-abstract reasoning. The motivation questionnaire examined interest, persistence, confidence, and value using age-appropriate statements. Module-validation and post-implementation forms assessed content, instructional quality, language, effectiveness, clarity, and usability.

Science-education and assessment reviewers examined item relevance, accuracy, curricular alignment, wording, response demand, and appropriateness for Grade 4. A group of 30 nonparticipant learners completed pilot measures. The source manuscript reported satisfactory internal consistency for the principal questionnaire (Cronbach's alpha = .85). Semi-structured interviews and open-response prompts were used to obtain explanations that could not be represented adequately by scale scores alone.

Procedure and Ethical Safeguards

After school authorization, parental or guardian consent and learner assent were obtained. Participation was voluntary, identifiers were excluded from analysis files, and withdrawal was permitted without academic penalty. Learners first completed the baseline measures. The validated module was then implemented through structured lessons and activities, followed by posttests, questionnaires, and selected interviews. Data were stored securely and reported in aggregate; short quotations were labeled with participant codes.

Data collection followed a consistent sequence. After authorization and consent procedures, learners completed the baseline conceptual-understanding and motivation measures. The classroom teacher then implemented the revised module through the planned lessons and activities. The posttest used the same construct coverage as the pretest, and learners completed the post-implementation motivation and feedback measures. Selected learners and teachers participated in short interviews about helpful features, confusing elements, resource constraints, and recommended changes.

The research involved minors. School authorization, written parental or guardian consent, learner assent, voluntary participation, confidentiality, and withdrawal without academic penalty were documented. Before submission, the author must insert the formal ethics approval details required for human-participant research: [INSERT ETHICS COMMITTEE NAME AND APPROVAL NUMBER].

Data Analysis

Frequencies, percentages, means, and standard deviations summarized needs, validation ratings, conceptual levels, motivation, and feedback. Paired-samples t tests examined pretest-posttest differences at alpha = .05, and Cohen's d quantified effect magnitude. Interview and open-ended responses were coded and grouped into recurring themes concerning engagement, understanding, clarity, usability, classroom constraints, and suggested improvements. Quantitative and qualitative findings were integrated during interpretation.

Descriptive statistics summarized ratings and distributions. Paired-samples t tests examined within-learner change, with alpha set at .05. Cohen's d described standardized mean differences (Cohen, 1988), and confidence intervals were reported when available. Because the study used the same learners at both measurement points, the paired analysis was appropriate; however, the manuscript did not provide complete assumption diagnostics or raw-score distributions, which should be retained for editorial review if requested.

Qualitative responses were analyzed thematically through repeated reading, initial coding, category development, comparison of learner and teacher accounts, and refinement of recurring patterns. This process followed the general logic of thematic analysis described by Braun and Clarke (2006). Short participant quotations were used to illustrate themes. Integration occurred during interpretation by examining whether stakeholder accounts explained, qualified, or challenged the quantitative pattern

3. RESULTS AND DISCUSSION

Research Question 1: What learner and teacher needs should inform the module?

The needs assessment showed a consistent but moderate pattern of conceptual difficulty. Learners reported the greatest difficulty with ecosystem interdependence (M = 2.46, SD = 1.18), followed by changes of state (M = 2.38, SD = 1.16), energy (M = 2.34, SD = 1.10), and forces (M = 2.10, SD = 1.05). The overall mean of 2.32 (SD = 1.12) indicates that difficulty was not confined to one topic; it involved explaining relationships and processes

ISSN (e): 3046-9570

across several Grade 4 competencies. Teacher accounts converged with these ratings and identified technical vocabulary, multi-step tasks, and limited opportunities for application as additional barriers.

Table 1. Selected Needs-Assessment Results

Indicator	Mean	SD	Interpretation
Difficulty understanding ecosystem interdependence	2.46	1.18	Somewhat agree
Difficulty explaining energy and how it works	2.34	1.10	Somewhat agree
Difficulty remembering types of forces	2.10	1.05	Somewhat agree
Difficulty understanding changes of state	2.38	1.16	Somewhat agree
Overall conceptual-difficulty mean	2.32	1.12	Somewhat agree

These findings support the decision to design the module around conceptual connections rather than additional factual coverage. Conceptual-change theory holds that learners need opportunities to expose prior ideas, compare them with evidence, and reorganize explanations (Posner et al., 1982). The reported difficulty with vocabulary and multi-step directions also indicates a cognitive-load problem: when language and procedures consume attention, fewer resources remain for integrating scientific relationships (Sweller et al., 2019). Accordingly, the module used short directions, familiar materials, guided inquiry, visual support, and sequenced prompts. This design is consistent with evidence that scaffolded inquiry is most productive when guidance helps learners connect activity with explanation (Hmelo-Silver et al., 2007).

Research Question 2: How was the module evaluated by teachers and science experts?

The 30 teachers and 20 science experts served as evaluators rather than intervention participants. Their ratings were positive across all domains: language appropriateness obtained the highest mean (M = 3.27), followed by instructional quality (M = 3.22) and content validity (M = 3.13). Reviewers particularly endorsed curricular alignment, logical sequencing, measurable objectives, real-life examples, and age-appropriate language. They nevertheless recommended stronger vocabulary support, more graphics, greater inclusivity, and differentiated guidance.

Table 2. Module Validation Summary

Domain	Mean	Interpretation
Content validity	3.13	Agree
Instructional quality	3.22	Agree to strongly agree
Language appropriateness	3.27	Strongly agree

The pattern suggests that the module was judged usable and instructionally coherent, but the comparatively lower content-validity mean signals room for substantive refinement. In the ADDIE model, validation is formative evidence used to improve alignment among objectives, learning activities, and assessment, not proof of learning effectiveness by itself (Branch, 2009). The recommendations concerning visuals and differentiation also qualify the high language rating: generally accessible wording does not eliminate the need to support learners with different reading proficiency and prior knowledge. The validation evidence therefore supports implementation on a provisional basis while identifying specific revisions that should accompany wider use.

Research Question 3: Did conceptual understanding change after module implementation?

The 50 learners who completed both assessments showed a clear shift in the structural complexity of their responses. At pretest, 48% were multistructural and 48% were relational, while none reached the extended-abstract level. At posttest, the multistructural proportion declined to 14%, 46% were relational, and 30% reached the extended-abstract level. Five learners (10%) were classified as unistructural after implementation, showing that improvement was not uniform and that some learners continued to require targeted support.

Table 3. Learners' SOLO Levels Before and After the Intervention

SOLO level	Pretest, n (%)	Posttest, n (%)
Unistructural	2 (4%)	5 (10%)
Multistructural	24 (48%)	7 (14%)

ISSN (e): 3046-9570

SOLO level	Pretest, n (%)	Posttest, n (%)
Relational	24 (48%)	23 (46%)
Extended abstract	0 (0%)	15 (30%)
Total	50 (100%)	50 (100%)

The paired comparison confirmed a statistically significant within-learner gain. The pretest-minus-posttest mean difference was -0.52 SOLO-level units, $t(49) = -5.987$, $p < .001$, 95% CI [-0.69, -0.35], with a large standardized effect ($d = 0.85$). The first null hypothesis was therefore rejected.

Table 4. Paired Test for Conceptual Understanding

Comparison	Mean difference	SD	t	df	p	95% CI	Cohen's d
Pretest - posttest	-0.52	0.614	-5.987	49	< .001	[-0.69, -0.35]	0.85

The distributional change is educationally important because it reflects movement from listing several relevant ideas toward connecting and extending them, which is the progression described by the SOLO taxonomy (Biggs & Collis, 1982). The module required learners to observe familiar materials, explain evidence, compare cases, and apply principles in new situations; these practices are consistent with the National Research Council's (2012) account of science learning as coordinated engagement with disciplinary ideas and scientific practices. The result also accords with research showing that active learning can improve performance when learners must reason with content rather than merely receive information (Freeman et al., 2014). However, the one-group design does not isolate the module from testing effects, maturation, ordinary instruction, or teacher influence. The effect size therefore describes change during the implementation period, not a definitive causal effect.

Research Question 4: Did science-learning motivation change after module implementation?

Motivation among the same 50 learners increased from a low pre-intervention mean of 1.78 to a high post-intervention mean of 3.23, representing a gain of 1.45 points on the four-point scale. The paired test was statistically significant, $t(49) = 12.83$, $p < .001$, and the reported standardized effect was very large ($d = 1.81$). The second null hypothesis was therefore rejected.

Table 5. Motivation Before and After Module Use

Measure	Result	Interpretation
Pre-intervention mean	1.78	Low
Post-intervention mean	3.23	High
Mean gain	1.45	Improved motivation
Paired test	$t(49) = 12.83$, $p < .001$	Statistically significant
Effect size	$d = 1.81$	Very large

The increase is consistent with self-determination theory. Progressive tasks and feedback may have supported competence, contextualized choices and reflection may have supported autonomy, and paired or group activities may have supported relatedness (Ryan & Deci, 2020). Learner accounts further indicate that interest was connected to understanding and usefulness rather than novelty alone. This interpretation is compatible with multidimensional views of science motivation that include intrinsic interest, confidence, self-determination, and perceived relevance (Glynn et al., 2011), as well as evidence that carefully designed science environments can connect engagement with substantive learning (Schneider et al., 2020). Nonetheless, self-report scores may be affected by social desirability, recent experience, and attention associated with a new intervention. A delayed measure and behavioral indicators are needed to determine whether the motivational increase was sustained.

Research Question 5: How did learners and teachers evaluate effectiveness, clarity, usability, and needed improvements?

Post-implementation feedback was gathered from the 50 learners who used the module and the 30 teachers who evaluated its classroom application (combined $n = 80$). Their aggregated ratings were strongly positive. Clarity received the highest overall rating ($M = 3.39$, $SD = 0.74$), followed by usability ($M = 3.38$, $SD = 0.72$) and effectiveness ($M = 3.32$, $SD = 0.78$). Because these are combined ratings from respondents with different roles, they should be interpreted as an overall appraisal rather than as separate learner and teacher estimates.

Table 6. Aggregated Learner and Teacher Post-Implementation Feedback

ISSN (e): 3046-9570

Domain or indicator	Mean	SD	Interpretation
Effectiveness			
Helped learners understand science concepts	3.40	0.70	Strongly agree
Lessons were useful and relevant to real life	3.30	0.79	Strongly agree
Improved interest and motivation	3.26	0.85	Strongly agree
Overall effectiveness	3.32	0.78	Strongly agree
Clarity			
Instructions were easy to understand	3.40	0.76	Strongly agree
Examples and explanations were clear and simple	3.44	0.67	Strongly agree
Activities were explained in an easy-to-follow manner	3.32	0.79	Strongly agree
Overall clarity	3.39	0.74	Strongly agree
Usability			
Module was well organized and easy to use	3.42	0.67	Strongly agree
Layout, images, and design supported engagement	3.34	0.77	Strongly agree
Time required for activities was reasonable	3.38	0.73	Strongly agree
Overall usability	3.38	0.72	Strongly agree

The qualitative evidence both explained and qualified the ratings. Learners described five recurring experiences: enjoyment of hands-on activities, improved conceptual understanding, clearer connections with daily life, increased confidence and curiosity, and a need for simpler vocabulary and more visuals. One learner stated, “It helped us understand better with experiments and real-life examples” (L4). Teachers likewise reported that the module supported the learning objectives, motivation, and classroom adaptation, but they emphasized the need for scaffolding, manageable activity length, low-cost materials, and alternatives where technology was unavailable. As one teacher observed, “It was highly usable, but adjustments were needed due to limited resources and lack of technology” (T3).

The convergence of strong ratings and specific reservations is more informative than either source alone. The ratings establish broad acceptability, while the interviews show the conditions under which that acceptability can translate into classroom use. Clear materials still required teacher questioning and modeling; engaging tasks still required time and resources; and contextual examples still required explicit links to scientific principles. This finding supports treating the module as a flexible instructional resource rather than a self-sufficient program. It also reflects the principle that scaffolding should enable increasingly independent reasoning while remaining responsive to learner needs (Hmelo-Silver et al., 2007).

Integrated Interpretation

Across the five research questions, the evidence forms a coherent but appropriately bounded account. Diagnosed conceptual and motivational needs informed the module design; teachers and experts judged the resulting material acceptable; the 50 implementation learners demonstrated significant within-group gains in conceptual understanding and motivation; and learner and teacher feedback identified plausible mechanisms as well as practical constraints. The findings therefore support the module's promise, clarity, and feasibility in the study setting. They do not establish that the module alone caused the gains or that the same effects will occur in other schools. A multi-site comparison study with implementation-fidelity measures, delayed posttests, disaggregated feedback, and complete assumption diagnostics is required before making claims about comparative effectiveness or broader adoption.

4. CONCLUSION

The study found that the contextualized Grade 4 Science module was responsive to identified learner needs and was positively evaluated by teachers and science experts for content, instructional quality, and language. Among the 50 learners who completed the main implementation, conceptual understanding increased significantly by 0.52 SOLO-level units, with movement toward relational and extended-abstract responses, and motivation increased from M = 1.78 to M = 3.23. Aggregated learner and teacher feedback also rated the module highly for effectiveness, clarity, and usability, while qualitative accounts identified hands-on inquiry, scaffolding, collaboration, and familiar real-life examples as helpful features. These findings support use of the module as a supplementary and adaptable resource in the study setting, provided that vocabulary, visuals, pacing,

differentiation, and low-cost activity options are further refined. Because the evidence came from one school, convenience sampling, self-report measures, and a one-group pretest-posttest design, the findings demonstrate promising within-group change and feasibility but do not establish exclusive causation, comparative superiority, or generalizability to other populations.

Limitations and Recommendations

The principal limitations are the one-group design, convenience sampling, one-school setting, modest learner sample, and reliance on some self-reported outcomes. Reuse of related measures may also have introduced testing effects. Future studies should employ a comparison group, random or matched allocation where feasible, delayed posttests, multi-school samples, and fidelity measures. Reporting should also include instrument-level reliability coefficients, complete score distributions, assumption checks, and disaggregated results where ethically and statistically appropriate.

For classroom use, the module should be refined through simpler technical vocabulary, additional diagrams and guided examples, differentiated tasks, low-cost experiment options, and clearer project rubrics. Teacher orientation should accompany broader implementation so that activities can be adapted without weakening their conceptual progression. Environmental applications such as waste segregation, biodegradability, and responsible material use may be retained because they connect scientific reasoning with everyday civic responsibility.

5. CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest related to the research or publication of this article.

6. USE OF GENERATIVE AI

The authors declare that no generative artificial intelligence (Generative AI) tools were used in the preparation, analysis, or writing of this manuscript.

7. REFERENCES

Biggs, J. B., & Collis, K. F. (1982). Evaluating the quality of learning: The SOLO taxonomy. Academic Press.

Branch, R. M. (2009). Instructional design: The ADDIE approach. Springer. https://doi.org/10.1007/978-0-387-09506-6

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2011). Science Motivation Questionnaire II: Validation with science majors and nonscience majors. Journal of Research in Science Teaching, 48(10), 1159–1176. https://doi.org/10.1002/tea.20442

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning. Educational Psychologist, 42(2), 99–107. https://doi.org/10.1080/00461520701263368

National Research Council. (2012). A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. National Academies Press. https://doi.org/10.17226/13165

OECD. (2023). PISA 2022 assessment and analytical framework. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66(2), 211–227. https://doi.org/10.1002/sce.3730660207

Prince, M. (2004). Does active learning work? A review of the research. Journal of Engineering Education, 93(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective. Contemporary Educational Psychology, 61, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Schneider, B., Krajcik, J., Lavonen, J., & Salmela-Aro, K. (2020). Learning science: The value of crafting engagement in science environments. Yale University Press.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, 31, 261–292. https://doi.org/10.1007/s10648-019-09465-5