American Journal of Humanities and Social Sciences Research (AJHSSR)
e-ISSN : 2378-703X
Volume-10, Issue-11, pp-134-146
www.ajhssr.com
Research Paper Open Access

Improving Grade VIII Students' Performance in Numerical Extended-Response Questions in Science Through the UICSC Problem-Solving Model: An Action Research Study

Pema Tshering[1], Deki Choden[2], Sangay Wangmo[3], Zangmo[4], Norbu Tshering[5]
(Gaupel Lower Secondary School, Paro, Bhutan)

ABSTRACT: Students often perform well on recall-based science items but struggle with Numerical Extended-Response Questions (NERQs), which require interpreting a problem, identifying given and required quantities, selecting a formula, showing calculations, and stating answers with correct units. This classroom-based action research examined whether the UICSC (Understand, Identify, Choose, Solve, Check) problem-solving method could improve Grade VIII students' NERQ performance at Gaupel Lower Secondary School, Paro, Bhutan. A single-group pre-test/post-test design was used, with a final paired dataset of 48 students. We analysed scores using a paired-samples t-test and a Wilcoxon signed-rank test. The class mean rose from 11.30/20 (56.5%) on the pre-test to 14.08/20 (70.4%) on the post-test, a mean gain of 2.78 marks (95% CI [2.02, 3.55]). The difference was statistically significant, $t(47) = 7.32$, $p < .001$, with a large within-group effect size (Cohen's $dz = 1.06$). The Wilcoxon test gave a consistent result ($W = 26.50$, $p < .001$). The number of students scoring below 8 fell from 11 to 1. Students' NERQ performance improved during the intervention period, suggesting that UICSC is a promising approach to structured problem solving. However, the single-group design, which has no comparison group, does not establish that the intervention alone caused the improvement.

KEYWORDS : *Formative assessment, instructional scaffolding, numerical extended-response questions, science education, UICSC problem-solving model.*

I. INTRODUCTION

Assessment is a fundamental element of teaching and learning. It involves the systematic gathering, interpretation, and use of evidence to understand what learners know, understand, and can do (CCEA, 2013; Linn & Gronlund, 2000). Assessment serves purposes beyond measuring achievement; it can inform instructional decisions, identify learning difficulties, and guide improvements in teaching and learning (Broadfoot et al., 2002). Assessments are commonly categorised by purpose as diagnostic (identifying prior knowledge and learning needs), formative (monitoring progress and improving learning), or summative (evaluating achievement at the end of a learning period). Assessment is formative when evidence of learning is used to modify teaching and learning (Black & Wiliam, 1998) and when learners understand expectations, identify gaps in their performance, and take steps to address them (Sadler, 1989).

In science education, assessment is particularly important because students are expected to do more than recall scientific facts. They must also demonstrate conceptual understanding, scientific reasoning, and the ability to apply knowledge to solve problems. Different assessment formats provide opportunities to assess these competencies. Objective-type questions assess students' ability to recall and recognise scientific concepts, whereas extended-response questions require learners to use their knowledge to explain phenomena or solve problems rather than simply recall information (Vo & Simmie, 2025). Numerical Extended-Response Questions (NERQs) are constructed-response items that require students to demonstrate understanding by solving numerical problems and showing the reasoning and calculations that lead to an answer (Livingston, 2009). Unlike multiple-choice questions, NERQs require students to generate their own answers and explain how they arrived at them. To solve NERQs successfully, students need to interpret the problem, identify the given and required quantities, choose an appropriate formula, perform calculations, and express the final answer with the correct units. Students may understand the relevant science but still have difficulty applying that knowledge systematically. Assessing their written solutions can therefore help teachers identify where difficulties occur and provide targeted feedback to improve learning (Black & Wiliam, 1998; Hattie & Timperley, 2007).

Research on problem-solving in science, particularly physics, has found that students' difficulties with quantitative problems often arise from the problem-solving process itself rather than simply from a lack of conceptual knowledge (Docktor & Mestre, 2014). Larkin et al. (1980) and Chi et al. (1981) also found that novices tend to select equations based on surface features, whereas experts first interpret the situation and identify the underlying principles. Pólya (1945) proposed four stages of problem-solving: understanding the problem, devising a plan, carrying out the plan, and looking back. Cognitive Load Theory explains why worked examples can assist novice learners by making solution procedures visible and reducing demands on working memory (Sweller, 1988; Sweller & Cooper, 1985). Scaffolding provides temporary support that is gradually withdrawn as learners become more independent (Wood et al., 1976). These approaches provide a basis for teaching students to interpret numerical problems, select appropriate formulas, show intermediate calculation steps, and check their solutions.

A key aim of integrated STEM education is to apply scientific and mathematical knowledge to solve problems (Kelley & Knowles, 2016); however, international and regional evidence suggests that applying such knowledge can be challenging. Turşucu et al. (2020) studied six Grade 10 students solving algebraic physics problems and found that students had difficulty applying algebraic skills and often used ad hoc procedures rather than systematic reasoning. Abdulla and Mumthas (2023) administered a 70-item physicomathematical conceptual test to 880 higher-secondary students in Kerala, India. Their findings highlighted difficulties in recognising formulas, extracting information from diagrams and graphs, and applying mathematics to physics problems. Jamila et al. (2024) investigated the effect of Student Teams–Achievement Divisions (STAD) on physics numerical problem-solving among 30 secondary students in Pakistan. Using a quasi-experimental design with two groups, they found that students taught through STAD improved their problem-solving performance compared with those taught through traditional methods. Although these studies document difficulties in numerical problem-solving and the potential benefits of structured instruction, they mainly concern older students and do not establish how a concise problem-solving routine might support Grade VIII science learners in Bhutan.

In Bhutan, the Programme for International Student Assessment for Development (PISA-D) provides national evidence concerning students' application of knowledge. The Bhutan Council for School Examinations and Assessment (BCSEA, 2019) reported challenges in tasks requiring students to formulate situations mathematically and interpret information in scientific contexts. These findings emphasise the importance of helping learners move beyond rote recall towards applying knowledge, interpreting information, and solving unfamiliar problems. Tenzin (2019) studied Grade IX students at Drukgyel Higher Secondary School using classroom observations, class tests, and student questionnaires. The study presented quantitative findings through bar graphs and pie charts. Tenzin reported that most students performed well following the intervention and recommended modelling different problem-solving methods and providing guided practice with numerical questions. As the findings were presented descriptively and came from a single school, they indicate promise rather than establish effectiveness. Nevertheless, the study is particularly relevant because it addresses numerical problem-solving in the Bhutanese school context. The present study extends this work by focusing on Grade VIII students, examining their written solution processes, and investigating a clearly sequenced UICSC (Understand, Identify, Choose, Solve, Check) routine in a lower-secondary science classroom.

The researcher analysed the 2026 Grade VIII Science midterm examination results at Gaupel Lower Secondary School, Paro Dzongkhag, and identified a difference in students' performance between objective-type questions and NERQs. In the objective-type section, 69.6% of students scored at or above the threshold of 27.9 out of 40 marks, compared with 51.02% in the NERQ section, a difference of 18.58 percentage points. These percentages are descriptive and should not be interpreted as a like-for-like comparison because the two sections may differ in total marks, content, and difficulty. The teacher's initial review of students' written responses also suggested that some students did not consistently identify the given and required quantities, show calculation steps, or include appropriate units. These procedural difficulties point to a need to support students in organising information and demonstrating the reasoning required to solve numerical problems. The findings are based on the school's 2026 midterm assessment and classroom review of students' written work; they represent local classroom evidence rather than nationally representative findings.

Although the assessment results indicate differences between the two question formats, objective-type and extended-response questions may differ in content, difficulty, cognitive demands, and marking criteria. Therefore, the observed performance difference alone cannot establish that the question format caused students' difficulties. Nevertheless, the written-script analysis identifies specific areas for instructional improvement: interpreting problems, identifying relevant quantities, selecting formulas, showing calculations, and using

appropriate units. The researcher's classroom observations also suggested that students could often recognise relevant scientific concepts but did not consistently demonstrate the procedural reasoning needed to solve numerical questions independently. Together, these findings support the need for a structured instructional approach that makes each stage of problem-solving explicit and provides opportunities for guided practice, feedback, and independent application.

Based on the issues identified above, the UICSC model, which builds on the problem-solving stages described by Pólya (1945), was selected as the intervention approach. Its five stages correspond to the difficulties observed in students' work and reflect established principles of problem-solving instruction. Understand requires students to interpret the problem and determine what it asks. Identify involves listing the given and required quantities. Choose guides students in selecting an appropriate formula. Solve requires students to substitute values and show the calculation steps. Check encourages students to verify units and assess whether an answer is reasonable. Like Pólya's (1945) framework, UICSC emphasises systematic reasoning, but it makes the procedures for numerical science questions more explicit. It also resembles mnemonic-based approaches such as RUCSAC, which guide learners through reading, understanding, choosing a method, solving, answering, and checking (Ehren et al., 2019). However, UICSC is organised specifically around identifying quantities and choosing an appropriate formula before solving, followed by checking the result. This explicit sequence is intended to address the procedural difficulties observed in the Grade VIII students' written responses.

The research gap can be understood in three related areas. First, although previous research has investigated numerical problem-solving, much of the evidence concerns higher-secondary or elder students, indicating a need for classroom-based research focused specifically on lower-secondary learners. Second, Bhutanese research, including Tenzin (2019), provides a useful foundation, but further evidence is needed on structured approaches to numerical extended-response questions among Grade VIII science students. Third, there is a need to examine whether a short, explicit method can address identifiable procedural errors in students' written solutions while supporting the application of scientific knowledge. Recent studies from India and Pakistan document difficulties in integrating mathematical concepts with physics (Abdulla & Mumthas, 2023) and show that a structured cooperative-learning approach can improve problem-solving performance (Jamila et al., 2024). However, these studies do not directly establish the effectiveness of UICSC for Bhutanese lower-secondary science learners.

This action research, titled 'Improving Grade VIII Students' Performance in Numerical Extended-Response Questions in Science Through the UICSC Problem-Solving Model: An Action Research Study', therefore investigates the use of UICSC to improve students' performance in NERQs at Gaupel Lower Secondary School. The intervention integrates teacher modelling, guided practice, scaffolding, formative feedback, and opportunities for peer and self-assessment. Students' learning is examined through pre-test and post-test assessments, supported by classroom observations of their problem-solving processes. By focusing on a specific lower-secondary classroom and the procedural difficulties evident in students' written work, the study aims to generate practical evidence on the use of a structured problem-solving routine to strengthen students' ability to apply scientific knowledge and communicate numerical solutions systematically.

II. LITERATURE REVIEW

Students in Grade VIII Science may demonstrate relatively strong performance in objective-type assessments, such as multiple-choice, fill-in-the-blank, and true/false items. These formats can assess recall and recognition, whereas NERQs require learners to interpret a problem, apply a formula, calculate, and present an answer with the correct unit. The literature reviewed below examines the conceptual and procedural demands of these tasks and instructional approaches that may support learners.

A NERQ is best understood as a sequence of linked steps rather than a single skill. Pólya (1945) described problem-solving as understanding the problem, devising a plan, carrying out the plan, and looking back at the result. Newman (1977) proposed a similar hierarchy for written problems, in which errors may occur at the stages of reading, comprehension, transformation (choosing the operation or formula), process skills (carrying out the calculation), and encoding (writing the answer in an appropriate form). Applied to science, these stages correspond closely to the demands of the present study: understanding the question, selecting and applying the formula, calculating accurately, and stating the answer with the correct unit. A learner who encounters difficulty at an early stage may struggle with subsequent stages, which helps explain why a student can perform well on objective items yet score poorly on a NERQ.

Research on expert and novice problem-solving helps explain why these difficulties occur. Chi, Feltovich, and Glaser (1981) found that novices tend to classify physics problems by surface features, whereas experts classify them according to underlying principles. In a Grade VIII context, this means that students may select a formula because the numbers or keywords in a question look familiar rather than because they understand which quantity is being asked for and how the quantities are related. When formulas are memorised without a clear understanding of what each symbol and unit represents, students may also find it difficult to rearrange formulas, convert units before substitution, or judge whether a final answer and its unit are reasonable.

Sweller's (1988) Cognitive Load Theory offers a further explanation. Conventional problem-solving can impose a heavy cognitive load on novice learners, who must hold the goal, given information, formula, and arithmetic in working memory simultaneously, leaving limited capacity for learning. Atkinson et al. (2000) further explained that studying worked examples before attempting similar problems helps novices build the schemas they need. Worked examples are among the well-supported instructional approaches in this area. For teachers, this implies that numerical problem-solving should be taught explicitly through modelled solutions and a clear routine rather than being assumed to develop through repeated testing.

Taken together, the literature suggests that poor performance in NERQs is not simply a result of insufficient effort or limited formula knowledge. It may reflect gaps in understanding the problem, connecting concepts with formulas, organising written solutions, and checking units and answers. Addressing these difficulties therefore requires strategies that build conceptual understanding, teach a clear problem-solving routine, and provide regular feedback on the stages at which students lose marks. Constructivist theory, particularly inquiry-based learning (IBL), is relevant because it promotes the active construction of knowledge rather than rote memorisation (National Research Council, 2000; Minner et al., 2010).

2.1 Inquiry-Based Learning (IBL) and Conceptual Understanding of Formulas

Inquiry-based learning (IBL) is widely recognised as an approach to developing scientific reasoning and higher-order thinking by engaging students in questioning, investigating, and constructing explanations. Kang (2022) reports positive effects on science literacy, while Dah et al. (2024) report gains in critical thinking and problem-solving. These findings suggest that IBL can help students develop a conceptual understanding of scientific principles and use their knowledge in a structured way to solve NERQs. For numerical topics, IBL can give formulas meaning. When students measure, vary, and compare quantities (for example, distance and time, or force and area) and investigate the relationships between them, a formula becomes a summary of something they have experienced. Units also become part of the meaning of a quantity rather than an afterthought. However, IBL can require substantial instructional time and teacher expertise, which may limit its consistent use. Inquiry alone also does not guarantee that students can present a correct written solution.

2.2 Structured Problem-Solving Strategies, Worked Examples and Scaffolding

Because conceptual understanding does not automatically produce a well-organised solution, students benefit from an explicit routine that makes each stage visible: identifying the given data and unknown, converting units where necessary, writing the formula, substituting values with units, calculating, and stating and checking the final answer. Such routines are consistent with Pólya's (1945) phases and Newman's (1977) stages. They can be introduced through worked examples (Atkinson et al., 2000; Sweller, 1988), followed by partially completed and then independent problems. McLure (2023) found that structured supports help learners who struggle with organisation, although that study concerned written explanations. Similar supports, such as solution templates and checklists, can be adapted for numerical solutions. If scaffolds are not gradually withdrawn, however, they may lead to learner dependence.

These strategies can be supported by higher-order questioning. Tagnin and Ní Ríordáin (2021) suggest that analytical questioning stimulates deeper classroom discussion, while Yanti and Thohir (2024) identify questioning as a way to develop reasoning skills. In numerical problem-solving, this means asking not only 'What is the answer?' but also 'What information is given and what is being asked?', 'Why is this formula appropriate for the problem?', and 'What does the unit of your answer tell you?' Teachers may, however, default to recall-based questioning because of curriculum pressures.

2.3 Assessment as Learning

Formative assessment and feedback are equally important. Yan and Pastore (2022) found that formative practices, including self-assessment and classroom discussion, are positively related to achievement and can help identify misconceptions. Hattie and Timperley (2007) showed that feedback is most effective when

it tells learners where they are in relation to the goal, what they need to do next, and how to improve the process they used rather than simply indicating whether an answer is correct. For numerical ERQs, this suggests marking and providing feedback at each stage (understanding, formula, calculation, and unit) rather than focusing only on the final answer. Analysing errors through a framework such as Newman's can also help teachers diagnose the main difficulties in a class. Constructed-response tasks provide more in-depth evidence of understanding than recall tasks (Vo & Simmie, 2025), and assessment tasks should promote higher-order thinking (Suhirman et al., 2022). Peer and self-assessment using a simple checklist can help learners understand what constitutes a complete solution. However, training is needed to ensure reliable peer feedback, and formative assessment can increase teachers' workload.

The literature indicated that success in numerical ERQs depends on three connected elements: conceptual understanding of the quantities and formulas involved; a clearly and explicitly taught problem-solving routine supported by worked examples and gradually withdrawn scaffolds; and formative assessment that provides stage-by-stage feedback. Much of the existing scholarship examines these strategies in isolation, while much of the ERQ literature concerns explanatory rather than calculation-based responses. There is limited empirical evidence on how an integrated approach works in a classroom, particularly in developing-country settings where resources, teacher workload, and assessment practices may constrain innovation.

Therefore, this study seeks to address this gap by implementing an integrated instructional and assessment approach that combines concept-first inquiry, a structured problem-solving routine with worked examples, higher-order questioning, and formative feedback through peer and self-assessment. The aim is to improve students' performance in numerical ERQs and provide classroom-based evidence and a practical framework for teachers who wish to help students understand questions, apply formulas, calculate accurately, and communicate answers with the correct units.

2.4 Main Action Research Question

How can the UICSC problem-solving model be used to improve Grade VIII students' performance in numerical extended-response questions in science?

2.5 Action Research Questions

1. What difficulties do Grade VIII students face when answering numerical extended-response questions in science?
2. How can the UICSC steps be integrated into classroom teaching of numerical problem-solving?
3. To what extent does the UICSC model improve students' performance in these questions?

2.6 Research objectives

The objectives of this action research are:

1. To identify the difficulties that Class VIII students face when answering NERQs in Science.
2. To implement the UICSC instructional cycle in Class VIII Science.
3. To examine changes in students' overall NERQ scores after the intervention.

III. RESEARCH METHODOLOGY

3.1 Research Setting and Participants

The study was conducted in the Grade VIII Science class at Gaupel Lower Secondary School, Paro, Bhutan. All 48 students enrolled in the class (20 boys and 28 girls) were included in the study. The intact class was selected because the researcher had identified a learning gap in students' performance on numerical extended-response questions. The study therefore used purposive whole-class participation within the research setting rather than random sampling.

3.2 Research Design

NERQs are a cornerstone of science assessment, as they require students to demonstrate conceptual understanding, select and apply formulas, calculate accurately, and communicate answers with correct units. Objective-type questions are mostly testing recall. Numerical ERQs are different in that they require learners to work through a problem in several connected steps. Despite performing well in objective-type assessments, many students experience significant difficulties when solving and presenting numerical problems in full.

To address this problem, this action research used an exploratory sequential mixed-methods design to gain a thorough understanding of the factors that contribute to students' difficulties in numerical ERQs and to evaluate the effectiveness of instructional and assessment interventions (Creswell & Plano Clark, 2018; Ivankova & Wingo, 2018)

The mixed-methods approach was chosen because it helps in bringing together quantitative and qualitative data, leading to a more complete understanding of the research problem. Quantitative data was collected before and after the intervention, to quantify the impact of the intervention on students' performance on numerical ERQs. This enabled the researcher to assess the extent to which the intervention improves students' achievement.

Qualitative data were collected through classroom observations and analysis of students' written test. The methods offered insight into the nature of students' difficulties (e.g. whether marks are lost in understanding the question, choosing the formula, calculating or writing units), their approach to problem solving and engagement with the intervention. The combination of quantitative and qualitative findings provides a deeper understanding of the challenges students face and inform strategies for improving their performance in numerical ERQs.

3.3 Research Tools

As this study employed an exploratory sequential mixed-methods approach, both quantitative and qualitative data collection methods were utilised. Quantitative data were collected through a pre-test and a parallel post-test of Numerical Extended-Response Questions (NERQs), each scored out of 20 using an analytic rubric. The rubric awarded marks for four criteria: (a) understanding the question (identifying the given and required quantities), (b) selecting and applying the correct formula, (c) accuracy of calculations, including unit conversion and substitution, and (d) providing the final answer with the correct unit. The test results provided evidence of improvement in students' overall achievement and performance by criterion, helping to evaluate the effectiveness of the intervention.

For the qualitative component, data were collected through classroom observations and analysis of errors in students' written solutions. Classroom observations enabled the researcher to examine students' engagement, participation, use of the problem-solving routine, and responses to the strategies implemented. Errors identified in students' scripts before and after the intervention were classified using Newman's (1977) stages, adapted for science-based numerical problem-solving, namely comprehension, formula selection, calculation, and units and answer presentation. The integration of quantitative and qualitative data provided a comprehensive understanding of students' difficulties in answering NERQs and the effectiveness of the intervention in addressing them.

3.4 Sampling Method

This study employed a purposive whole-class sampling technique, involving all students in Class VIII. The students were selected because they had demonstrated difficulties in answering Numerical Extended-Response Questions (NERQs). Whole-class participation was appropriate for this action research because the intervention aimed to address an identified learning difficulty within the researcher's own classroom. This sampling strategy enabled the researcher to gather relevant evidence about students' difficulties in solving numerical problems and to gain meaningful insights into the challenges they encountered.

3.5 Data collection and analysis method

The data analysis was carried out in phases. For the quantitative phase, a paired-samples t-test was conducted to compare students' pre-test and post-test scores in NERQs to determine whether there was a statistically significant difference in their performance following the intervention. The significance level was set at 0.05. In addition, the mean scores for each rubric criterion (understanding the question, formula, calculation, and unit) were compared between the pre-test and post-test to identify the areas in which students demonstrated the greatest improvement.

For the qualitative phase, thematic analysis was employed to analyse the data collected through classroom observations. The observational data were systematically examined to identify, code, and categorise recurring patterns and themes related to students' approaches, participation, and difficulties in solving NERQs. This process facilitated a comprehensive interpretation of classroom behaviours and provided deeper insights into the factors that influenced students' performance during the intervention.

3.6 Ethical Considerations

Permission to conduct the study was obtained from the school authority, and written consent was obtained from parents/guardians of the participating students. Students were informed about the purpose of the study and their right to withdraw without academic consequences. The research team took care to ensure that participation or non-participation did not affect students' academic standing.

Confidentiality was maintained by using anonymised participant codes in the analysis and presentation of results rather than students' names. Assessment data were used for research purposes and stored securely. The figures in this manuscript do not display student names or roll numbers.

The researcher ensured that data were handled, analysed, and reported honestly and objectively. All research information was stored securely.

3.7 Significance of the Study

The present study is important because it seeks to address the persistent gap between students' performance in objective-type questions and NERQs in Class VIII Science. It identifies the specific difficulties students experience when solving numerical problems and provides insight into their learning needs, particularly in understanding questions, selecting and using formulas, calculating accurately, and applying correct units.

The findings may help science teachers design more effective teaching and assessment strategies for numerical problems. In particular, stage-by-stage marking and feedback can help teachers move beyond focusing only on the final answer and identify precisely where students lose marks.

The study may help learners develop a more effective routine for answering NERQs by improving their ability to understand problems, use formulas, show their working, and write answers with the correct units. This may contribute to improved performance in numerical and other extended-response assessments.

The study may also provide useful insights for other subjects, such as Mathematics, where students encounter similar calculation-based problems. It contributes to classroom-based action research by demonstrating how instructional interventions can be used to strengthen students' problem-solving skills in science education.

IV. RESULTS

4.1 Pre-test result

The pre-test results showed a mean score of 11.30 out of 20 (56.5%) and a median of 10.5, with scores ranging from 2.5 to 19.0 (*SD* = 4.26). This dispersion indicates considerable heterogeneity in students' baseline performance. Of the 48 students, 29 (60.4%) attained a score of 10 or above, whereas 19 (39.6%) scored below this threshold. Furthermore, 11 students (22.9%) scored below 8, while 10 students (20.8%) achieved scores of 16 or higher.

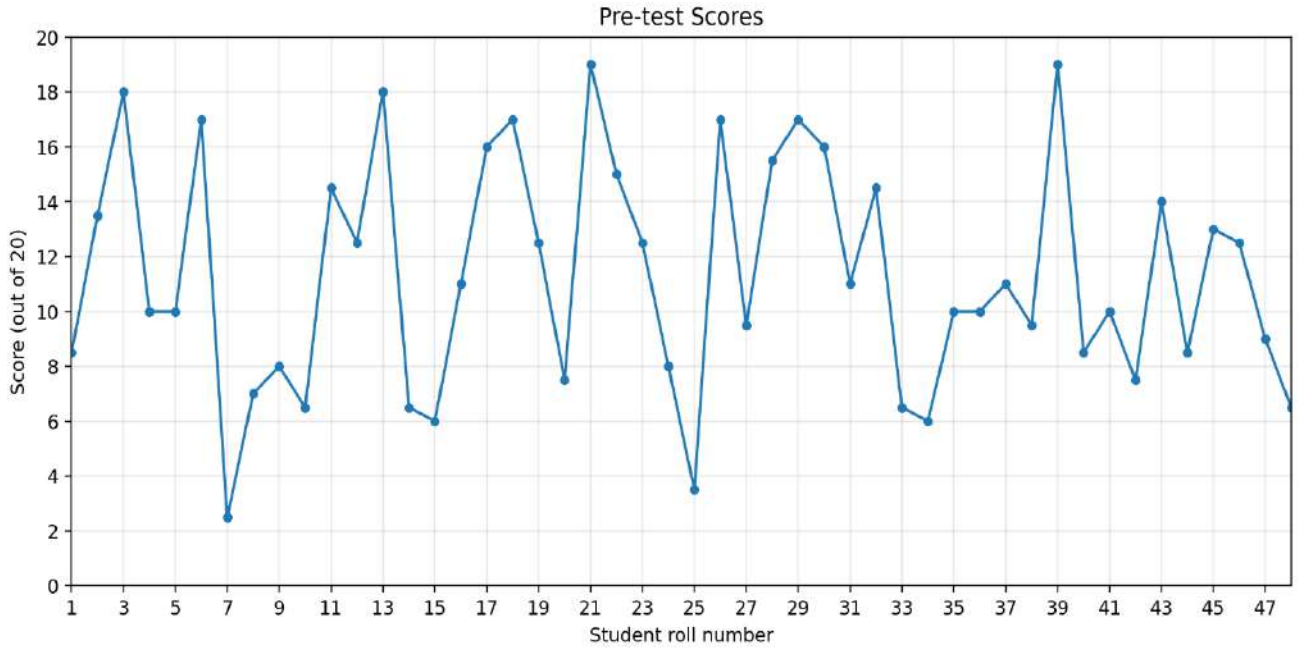

Fig 1. *Distribution of pre-test and post-test scores by score range (n = 48).*

The range of scores was 16.5 marks, which further illustrates the dispersion in baseline performance. Sixteen students (33.3%) scored between 8 and 11.5, while 11 (22.9%) scored between 12 and 15.5. These baseline groupings were used to examine descriptive changes in performance across starting levels. Overall, the pre-test results indicate heterogeneous baseline performance and a need for continued support in numerical problem-solving, particularly among students who began with lower scores.

The teacher's initial review of participants' written responses indicated difficulties in interpreting numerical questions, selecting and applying appropriate formulas, presenting calculation procedures, and including correct units. These observations informed the selection of the UICSC routine.

Analysis of students pre-test responses revealed that, although some students could identify relevant information in the questions, they had difficulty determining and applying the appropriate formula. Some attempted calculations without following a systematic approach, while others left questions unanswered because they could not identify the initial step. Students also experienced difficulty presenting their calculation procedures, reporting answers with appropriate units, and verifying the accuracy of their solutions. These observations underscored the need for structured scaffolding to support the development of systematic problem-solving skills.

4.2 Implementation of Intervention Strategies

Guided by the literature reviewed, this study adopted an integrated instructional and assessment framework rather than treating concept-first inquiry, structured problem-solving, worked examples, and formative assessment as separate interventions. The strategies were implemented with the Grade VIII Science class over one intervention cycle, from August to October 2026. The cycle comprised the following interrelated strategies:

4.3 Concept-first, inquiry-based introduction of formulas

Before students practised calculations, lessons focused on developing an understanding of the meaning of each formula through guided inquiry. Students measured or examined quantities (for example, distance and time, mass and volume, or force and area), investigated the relationships between them, and learned the symbols, formulas, and units together. This approach was intended to reduce rote memorisation and help students understand why units were necessary. It was informed by the work of Kang (2022) and Dah et al. (2024), which highlighted the importance of developing conceptual understanding of mathematical and scientific relationships.

4.4 A structured five-step problem-solving routine

Students were taught and practised a fixed five-step routine for numerical ERQs: (1) Understand: read the question and determine what is being asked; (2) Identify: list the given data and the unknown, using symbols and units; (3) Choose: select and write the appropriate formula, rearranging it if necessary; (4) Solve: check and convert units where necessary, substitute values, and calculate step by step; and (5) Check: write the final answer with the correct unit and assess whether it is reasonable. This sequence follows the broad principles of Pólya's (1945) framework and Newman's (1977) stages while making each stage of the solution visible and assessable.

4.5 Worked examples, scaffolding, and higher-order questioning

The teacher modelled worked examples aloud, then moved to partially completed problems and finally to independent tasks so that scaffolding could be gradually withdrawn (Atkinson et al., 2000; Sweller, 1988; McLure, 2023). 'Spot the error' tasks required students to identify and correct an incorrect formula, a missing unit conversion, or an omitted unit in a worked solution. Classroom questioning also shifted from recall-based questions (for example, 'What is the formula for…?') towards questions such as 'What information is given and what is being asked?', 'Why is this formula appropriate?', and 'What does the unit of your answer tell you?', drawing on Tagnin and Ní Ríordáin (2021) and Yanti and Thohir (2024).

4.6 Formative assessment, feedback, and peer/self-assessment.

Short, low-stakes NERQ tasks were used throughout the intervention cycle, in addition to the pre-test and post-test. Teacher feedback focused on the stages of a complete numerical solution: understanding the question, selecting a formula, showing calculations, and presenting an answer with the correct unit. Students used a checklist to review their own and peers' solutions and recorded recurring mistakes.

4.7 Alignment with the research tools.

The intervention sequence was aligned with the research tools. Overall pre-test and post-test scores provided the quantitative evidence, while classroom observations and the teacher's review of students' written responses provided descriptive context for interpreting the changes.

4.9 Post-intervention Data Collection

Following the intervention cycle, post-intervention data were collected using the same procedures and scoring instruments applied at baseline. This enabled comparison of students' performance before and after the intervention.

A post-test consisting of numerical extended-response questions was administered to the 48 participating students under normal classroom test conditions. The post-test was intended to be parallel to the pre-test in structure, topic coverage, and cognitive demand. Both assessments were scored out of 20, and the same overall scoring approach was used for the paired comparison. The available spreadsheet contains total scores only; criterion-level scores were not available for separate analysis.

4.11 Performance by pre-test score band

To examine whether the intervention benefited students with different starting levels, the 48 students were grouped according to their pre-test scores. Table 1 presents the number of students in each band and the mean pre-test, post-test, and gain scores.

Pre-test score band	Students (n)	Pre-test mean	Post-test mean	Mean gain
Below 8	11	6.00	11.32	+5.32
8 to 11.5	16	9.53	11.78	+2.25
12 to 15.5	11	13.64	16.14	+2.50
16 and above	10	17.40	18.55	+1.15

TABLE 1. Pre-test and post-test mean scores and gains by pre-test performance band (n = 48)

The largest mean gain was observed among students who scored below 8 in the pre-test (+5.32 marks). Students in the 8–11.5 band gained 2.25 marks, those in the 12–15.5 band gained 2.50 marks, and those scoring 16 or above gained 1.15 marks. These are descriptive subgroup comparisons; the smaller gain among the highest-scoring group may partly reflect a ceiling effect.

4.12 Quantitative and qualitative Evidence of students' progress

The post-test mean was 14.08 (SD = 3.82), with a median of 13.5 and scores ranging from 7.0 to 20.0. The number of students scoring below 8 decreased from 11 (22.9%) to 1 (2.1%), while the number scoring 10 or more increased from 29 (60.4%) to 40 (83.3%). Five students achieved full marks in the post-test. Forty-seven students improved their scores, while one student's score decreased.

Similarly, classroom observations during the implementation of the UICSC routine indicated that students became increasingly familiar with the five steps through teacher modelling, worked examples, collaborative activities, and the 'I do, We do, You do' approach. The teacher observed that students increasingly attempted to identify relevant information, select appropriate formulas, show calculation procedures, and check their answers. Some students continued to require additional practice and individual guidance. The classroom account also noted that students began applying the routine in Mathematics, suggesting possible transfer to other calculation-based learning tasks. These observations are descriptive and should not be interpreted as a separately measured perception outcome.



Fig 2. Paired pre-test and post-test scores for 48 students; individual lines are anonymised and the bold line shows the class mean.

The paired-score plot presents anonymised participant-level changes without displaying names or roll numbers. Forty-seven students recorded a higher post-test score than pre-test score, and one student recorded a lower score. Individual identifiers are not reported in the text to protect student confidentiality.

V. FINDINGS

This action research examines whether an integrated intervention built around the UICSC (Understand, Identify, Choose, Solve, Check) routine improves Grade VIII students' performance in Numerical Extended-Response Questions (NERQs). The discussion interprets the findings in relation to the baseline difficulties, the research questions, and the relevant literature.

5.1 Difficulties students face in answering NERQs

The teacher's initial review of students' written responses suggests difficulties in interpreting questions, selecting appropriate formulas, showing calculation steps, and presenting answers with appropriate units. The final dataset contains total pre-test and post-test scores only; therefore, criterion-level changes and the frequency of each error type cannot be calculated. The baseline total-score analysis indicates a mean of 11.30/20 (56.5%), with scores ranging from 2.5 to 19.0 (SD = 4.26).

5.2 Change in overall NERQ performance

Following the intervention, the class mean increases from 11.30/20 (56.5%) to 14.08/20 (70.4%), representing a gain of 2.78 marks. The paired-samples t-test indicates a statistically significant mean difference, $t(47) = 7.32$, $p < .001$, 95% CI [2.02, 3.55], with a large within-group effect size (Cohen's d_z = 1.06). The Wilcoxon signed-rank test also indicates a statistically significant increase ($W = 26.50$, $p < .001$).

The normalised gain is approximately 0.32, indicating that students close about 32% of the gap between their pre-test mean and the maximum possible score. This suggests a meaningful but incomplete improvement, with scope for further development. The normalised gain is interpreted descriptively and alongside the limitations of the single-group design.

The findings are consistent with earlier research indicating that structured instruction supports numerical problem-solving. Tenzin (2019) reported that Grade IX students in Bhutan performed well following an intervention that modelled problem-solving strategies and provided guided practice. Jamila et al. (2024) found that a structured approach outperformed traditional teaching in a study conducted in Pakistan. The present study adds classroom-based evidence of gains in total test scores among Grade VIII learners through the use of an explicit problem-solving routine. However, differences in context and study design require consideration when comparing these findings.

5.3 Factors that may contribute to the success of the intervention

Each UICSC stage addresses a difficulty identified in students' initial written responses. Understand encourages students to interpret the question; Identify prompts them to list the given and required quantities; Choose guides the selection of an appropriate formula; Solve supports unit conversion, substitution, and step-by-step calculation; and Check encourages students to review the final answer and its unit. Teacher modelling, worked examples, guided practice, and feedback may further support students in making these problem-solving stages explicit.

However, because the intervention combines several strategies and no criterion-level scores are available, the study cannot determine which stage or component contributes most to the observed score improvement. The study therefore examines the effectiveness of an integrated classroom approach rather than isolating the individual effects of the UICSC routine, worked examples, questioning, or feedback.

5.4 Limitations

This action research has several limitations that should be considered when interpreting the findings. First, the single-group pre-test/post-test design had no comparison group; therefore, the gains cannot be attributed to the intervention with certainty. Practice effects, normal learning over the study period, and increased familiarity with the test format may also have contributed. Second, although the post-test was designed to be parallel to the pre-test, equivalence in difficulty cannot be guaranteed. Third, the study involved one class of 48 students in one school, and the teacher was also the researcher, which may have introduced bias in scoring and observation. Fourth, the post-test was administered immediately after the intervention, so the study provides no evidence about the long-term retention of learning. Finally, the integrated design means that the contribution of each individual strategy cannot be isolated. The findings should therefore be regarded as promising classroom-level evidence rather than as results that can be generalised to a wider population.

VI. CONCLUSION

This action research was conducted to explore the problems faced by the Grade VIII students when answering the Numerical Extended-approach Questions (NERQs) in science and the application of UICSC (Understand, Identify, Choose, Solve, Check) problem-solving model as a pedagogical approach. The study was premised on the assumption that the ability to solve numerical problems is dependent on conceptual knowledge and systematic thinking rather than rote memorization of equations. The intervention provided students with a systematic pathway for interpreting and solving numerical issues through concept-first inquiry, instructor modelling, worked examples, scaffolding, higher-order questioning and formative feedback.

The results demonstrate that successful numerical problem solving cannot be explained solely by practice or formula retrieval. Students need to be explicitly taught how to carry out a succession of inter-related skills: comprehending the demands of a question; collecting relevant data; selecting appropriate formulae; carrying out calculation methods and communicating answers with proper units. The UICSC model offers a practical framework for developing these skills and fostering more organised, reflective and self-regulated learning behaviours, and the findings highlight the contribution of continuous formative assessment, constructive criticism, and differentiated support in catering for the needs of diverse learners.

There are some limitations that should be noted. These include the single classroom context, the small sample, and the lack of information as to whether gains were maintained over time. In future study, therefore, larger and more diverse samples, longitudinal designs, and comparison groups are required to determine the long-term effectiveness and transferability of the UICSC model. Nevertheless, this study highlights the possibility of reflective practice and evidence-informed teaching, and implies that structured problem-solving frameworks have tremendous promise for lower-secondary science and other calculation-based subjects. Finally, the objective of science education should be to transfer not only knowledge and formulas but also to build students' ability to think critically, logically and solve problems independently.

VII. RECOMMENDATIONS

Based on the findings and insights gained from this action research, the following recommendations are proposed to strengthen the teaching and learning of numerical extended-response questions (NERQs) in Science, support learners with diverse needs, and guide future research and classroom practices.

- **For science teachers:** Teach a problem-solving routine such as UICSC explicitly, model it through worked examples, and gradually withdraw support. Assess NERQs by stage (understanding, formula, calculation, and unit) so that students can identify where marks are lost.
- **For differentiated support**: Provide targeted assistance to students who continue to score below 10 and extension tasks for students who make limited gains because their initial scores are already close to the maximum.
- **For school practice:** Share the UICSC routine and checklist across Grade VIII Science classes and consider adapting them for Mathematics and other calculation-based subjects.
- **For future research:** Use a comparison group or a multi-class design, include a delayed post-test to examine retention, and interview students who make substantial gains to understand which aspects of the intervention they found most helpful. Further studies could also compare individual intervention components to identify their relative contributions.

Overall, the findings provide encouraging classroom-based evidence from Bhutan regarding the potential of an explicit, scaffolded problem-solving approach to improve lower-secondary students' performance in numerical extended-response questions. The study highlights the value of teaching problem-solving as a systematic process rather than focusing solely on formula recall and final answers.

VIII. ACKNOWLEDGEMENT

We extend our heartfelt appreciation to the School Administration of Gaupel Lower Secondary School for their continuous encouragement and moral support throughout this research.

We are particularly indebted to Mr. Sangay Tshewang, Training Director of the Technical Training Institute, Chumey, for his invaluable guidance, insightful feedback, and professional support as our critical friend.

We also sincerely acknowledge the students and colleagues for their cooperation and active participation, which contributed significantly to the successful completion of this study.

REFERENCES

[1] Council for the Curriculum, Examinations and Assessment. (2013). Assessment for learning: Guidance for key stages 1–3. CCEA.

[2] Linn, R. L., & Gronlund, N. E. (2000). Measurement and assessment in teaching (8th ed.). Prentice Hall.

[3] Broadfoot, P., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). Assessment for learning: 10 principles. Assessment Reform Group. https://hdl.handle.net/1893/32458

[4] Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. https://doi.org/10.1080/0969595980050102

[5] Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18, 119–144. https://doi.org/10.1007/BF00117714

[6] Vo, D. V., & Simmie, G. M. (2025). Assessing scientific inquiry: A systematic literature review of tasks, tools and techniques. International Journal of Science and Mathematics Education, 23(4), 871–906.

[7] Livingston, S. A. (2009). Constructed-response test questions: Why we use them; how we score them (R&D Connections, No. 11). Educational Testing Service. https://www.ets.org/research/policy_research_reports/publications/periodical/2009/hkap.html

[8] Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. https://doi.org/10.3102/003465430298487

[9] Docktor, J. L., & Mestre, J. P. (2014). Synthesis of discipline-based education research in physics. Physical Review Special Topics – Physics Education Research, 10(2), Article 020119. https://doi.org/10.1103/PhysRevSTPER.10.020119

[10] Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. Science, 208(4450), 1335–1342. https://doi.org/10.1126/science.208.4450.1335

[11] Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2

[12] Pólya, G. (1945). How to solve it. Princeton University Press.

[13] Sweller, J. (1988). Cognitive load during problem-solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

[14] Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem-solving in learning algebra. Cognition and Instruction, 2(1), 59–89. https://doi.org/10.1207/s1532690xci0201_3

[15] Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

[16] Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. International Journal of STEM Education, 3, Article 11. https://doi.org/10.1186/s40594-016-0046-z

[17] Turşucu, S., Spandaw, J., & de Vries, M. J. (2020). Search for symbol sense behavior: Students in upper secondary education solving algebraic physics problems. Research in Science Education, 50, 2131–2157. https://doi.org/10.1007/s11165-018-9766-z

[18] Abdulla, S. U., & Mumthas, N. S. (2023). Physico-mathematical conceptual difficulties: A barrier to learning 'motion' in physics among higher secondary students in Kerala. Journal of Indian Education, 48(4), 128–145.

[19] Jamila, Afzal, M. T., Mir, R., & Marium, A. (2024). Enhancing problem-solving skills of physics students through Student Team Achievement Division (STAD) at secondary level. Pakistan Journal of Society, Education and Language, 10(2), 245–254. https://pjsel.jehanf.com/index.php/journal/article/view/1401

[20] Bhutan Council for School Examinations and Assessment. (2019). Education in Bhutan: Findings from Bhutan's experience in PISA for Development. https://www.bcsea.bt/uploads/publications/Bhutan-PISA-D-National-Report-BCSEA-2019_1626326228.pdf

[21] Tenzin, D. (2019). Improvement of physics numerical problem-solving skills of students in Bhutan. Asian Journal of Education and Social Studies, 4(3), 1–13. https://doi.org/10.9734/ajess/2019/v4i330120

[22] Ehren, M., Wollaston, N., Goodwin, J., & Star, J. (2019). The nature, prevalence and effectiveness of strategies used to prepare pupils for Key Stage 2 maths tests: Final report. UCL Institute of Education. https://discovery.ucl.ac.uk/id/eprint/10085698/

[23] Newman, M. A. (1977). An analysis of sixth-grade pupils' errors on written mathematical tasks. Victorian Institute for Educational Research Bulletin, 39, 31–43.

[24] National Research Council. (2000). Inquiry and the national science education standards: A guide for teaching and learning. National Academies Press. https://doi.org/10.17226/9596

[25] Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—What is it and does it matter? Results from a research synthesis years 1984 to 2002. Journal of Research in Science Teaching, 47(4), 474–496. https://doi.org/10.1002/tea.20347

[26] Kang, J. (2022). Interrelationship between inquiry-based learning and instructional quality in predicting science literacy. Research in Science Education, 52(2), 339–355.

[27] Dah, N. M., Noor, M. S. A. M., Kamarudin, M. Z., & Azziz, S. S. S. A. (2024). The impacts of open inquiry on students' learning in science: A systematic literature review. Educational Research Review, 43, Article 100601.

[28] Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. Review of Educational Research, 70(2), 181–214. https://doi.org/10.3102/00346543070002181

[29] Tagnin, L., & Ní Ríordáin, M. (2021). Building science through questions in Content and Language Integrated Learning classrooms. International Journal of STEM Education, 8, Article 34.

[30] Yanti, F. A., & Thohir, M. A. (2024). Higher-order thinking skills in science learning: A systematic review from 2014–2023. International Journal of Evaluation and Research in Education, 13(4), 2488–2498.

[31] Yan, Z., & Pastore, S. (2022). Formative assessment practices in science education: A meta-synthesis study. Studies in Educational Evaluation, 75, Article 101186.

[32] Suhirman, S., Prayogi, S., & Asy'ari, M. (2022). Exploring the structure of students' scientific higher-order thinking in science education. Thinking Skills and Creativity, 43, Article 100999.

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

[34] Ivankova, N., & Wingo, N. (2018). Applying mixed methods in action research: Methodological potentials and advantages. American Behavioral Scientist, 62(7), 978–997. https://doi.org/10.1177/0002764218772673

[35] McLure, F. (2023). The Thinking Frames Approach: Improving high school students' written explanations of phenomena in science. Research in Science Education, 53(1), 173–191.