



SAJSE

SOUTHEAST ASIAN JOURNAL OF STEM EDUCATION

A publication of SEAMEO STEM-ED

ISSN: 2709-1252

Vol. 6 No. 3

July 2026



Southeast Asian Journal of STEM Education

A Publication of SEAMEO STEM-ED
132 Pages
ISSN: 2709-1252

2026
Vol. 6 No. 3
July 2026

Editors and Contributors

Editor-in-Chief:

Dr Orawan Sriboonruang

Editors:

Asst Prof Dr Burin Asavapibhop
Dr Norhaniza binti Abdul Hamid
Asst Prof Dr Nutthathirataa Withitwinyuchon

Reference Checking:

Veeraya Chowopass

Layout and Design:

Varangkana Ruengvisesh

SEAMEO STEM-ED

928, 11th Floor, Natural Science Building
Sukhumvit Road, Phra Khanong, Khlong Toei
Bangkok 10110, Thailand

Email: sajse@seameo-stemed.org
Website: <https://seameo-stemed.org/>

Printed in Thailand

© SEAMEO STEM-ED 2026

The designations employed and the presentation of the materials throughout the publication do not imply the expression of any opinion whatsoever on the part of SEAMEO STEM-ED concerning the legal status of any country, territory, city, or area, or of its authorities or frontiers or boundaries.

For submissions and inquiries please contact
Email: sajse@seameo-stemed.org



Contents

Page

PREFACE

i

EDITORIAL

ii

POLICY SECTION

OUT-OF-SCHOOL CHILDREN, YOUTH AND ADULT (OSCYA) LEARNERS STEM PERFORMANCE QUANTIFIER AND SENTIMENT: INSIGHT FROM ACCREDITATION AND EQUIVALENCY (A&E) TEST

2-14

Dionito F Mangao Jr

TEACHER SECTION

GAUTHMATH AS AI-ASSISTED SCAFFOLDING IN SECONDARY MATHEMATICS: LEARNER ACHIEVEMENT AND ACCEPTABILITY IN COMBINATORICS

16-28

Christian Genesis R Guinto and Maria Magdalena S Cabangca

STORIES TO EXPERIMENT: THE STREAM STORY CYCLE FOR COMMUNICATIVE ESL LEARNING

29-41

Hjh Nor Nasriah Akmal binti Haji Kasim

GUIDED INQUIRY IN PRIMARY SCIENCE LESSONS AND THE CONCURRENT DEVELOPMENT OF SCIENCE PROCESS SKILLS, 21ST CENTURY SKILLS, AND COMPUTATIONAL THINKING: A SYSTEMATIC LITERATURE REVIEW (2016-2026)

42-56

Tracy Mensan and Corrienna Abdul Talib

FACULTY ATTITUDES TOWARDS OUTCOME-BASED EDUCATION IMPLEMENTATION IN LAO PUBLIC UNIVERSITIES: A CASE STUDY OF ENGINEERING EDUCATION

57-70

Ayyia Machang, Yongxia Ma, and Houmlack Mingboubpha

NATURE OF ALGEBRA TASKS IN PHILIPPINE TEXTBOOKS FOR GRADES 9 AND 10: CONTENT TOPICS

71-86

Michael G Domanais and Ryan V Dio

DEVELOPING AND VALIDATING A MINECRAFT-BASED COMPUTATIONAL THINKING TEACHING MODEL FOR CODING EDUCATION IN PRIMARY SCHOOLS

87-100

Goh Kok Ming



Contents

Page

STUDENT SECTION

EFFICIENCY ASSESSMENT OF A FOOD WASTE SHREDDING MACHINE FOR BIOGAS PRODUCTION

102-108

Marie Antonette Anahaw, Ashly Claudet Grafilo, Harvey Oborza, and Cris Edison U Villanueva

DEVELOPMENT AND PERFORMANCE EVALUATION OF AN AI-INTEGRATED ULTRASONIC FLY REPELLENT SYSTEM WITH REAL-TIME HOUSE FLY DETECTION

109-121

*Kanyanat Khamhongsa, Premyuda Witchayateerakun, Thanaphon Phunamsiang, and
Chumpon Chareesaen*

PREFACE

Welcome to the Southeast Asian Journal of STEM Education (SAJSE), a prominent online platform established by SEAMEO STEM-ED, dedicated to the advancement of STEM education in Southeast Asia. As an open-access publication, SAJSE is an invaluable repository for insights, ideas, and innovations that define the future of STEM education in the region.

At SEAMEO STEM-ED, we believe in the power of collaboration and knowledge exchange to advance STEM education. The articles in SAJSE reflect this belief, offering a wealth of quality knowledge that transcends regional boundaries. Our dedication to openness is demonstrated by our usage of Creative Commons licencing, which allows for the free use, distribution, and replication of our published content with due credit.

We are excited to announce that we are opening up opportunities for teachers and students to share their voices and reflections from the classroom. Teachers are invited to share their innovative practices, challenges, and insights in teaching STEM subjects, providing valuable perspectives that can benefit educators across Southeast Asia. Similarly, students are encouraged to share their experiences and reflections on STEM science projects, highlighting the impact of hands-on learning experiences on their education and personal development.

This initiative aims to echo the voices of teachers and students, providing a platform for them to share their unique perspectives and contribute to the broader conversation on STEM education in Southeast Asia. By amplifying these voices, we hope to inspire others and drive positive change in STEM education practices and policies.

Authors who submit their papers to SAJSE will benefit from the journal's rigorous peer-review process, ensuring the quality and integrity of published articles. Additionally, published papers will be freely accessible to educators, researchers, and policymakers, facilitating the dissemination of innovative practices and ideas in STEM education.

We invite educators, researchers, and innovators from across Southeast Asia to join us in this quest to elevate STEM education. Together, let us construct a cooperative and inclusive learning ecosystem that equips the future generation of innovators and solution seekers, propelling Southeast Asia towards a future characterised by innovation and excellence.

Thank you for being a part of this journey.

Warm regards,

SEAMEO STEM-ED

EDITORIAL

Welcome to SAJSE, the leading platform for advancing STEM education and encouraging innovation in Southeast Asia. Our journal exemplifies the spirit of collaboration, creativity, and capacity building. We are dedicated to making STEM education accessible to all, while also offering chances for educators and students to share their research, teaching techniques, and new ideas. This open-access journal serves as a beacon for researchers, educators, policymakers, and students seeking excellence in STEM education.

A Platform for Diverse Voices

The Southeast Asian Journal on STEM Education transcends its role as a mere repository of information, establishing itself as a vibrant forum that embraces a plurality of perspectives. This journal curates a comprehensive array of contributions, encompassing seminal research articles from seasoned scholars, innovative pedagogical methodologies from educators, and introspective analyses on integrated STEM disciplines authored by students. It operates as a platform that acknowledges and values the diverse viewpoints of all constituents within the educational sphere, thereby facilitating a collective ethos of learning and developmental synergy.

Bridging Policy and Practice

A distinctive feature of this journal is its specialised policy section, designed to stimulate dialogue on the development and enactment of educational policies, with a particular emphasis on STEM disciplines. This section covers a broad spectrum, from theoretical propositions to empirical research and the exchange of best practices. The journal aims to contribute to and impact policy-making processes that determine the trajectory of educational advancement. It represents an effort to reconcile policy with practice, guaranteeing that educational innovations are firmly anchored in the practical landscape and meet the requirements of learners.

Global Perspectives with a Local Focus

Although the journal invites submissions from educators, policymakers, teachers, and students globally, it prioritises local implementations, research, or practices from Southeast Asian countries. This emphasis underscores a dedication to tackling the distinct challenges and seizing the opportunities present in the region, advocating for solutions that are informed by global perspectives yet tailored to local contexts.

A Call to Action

We warmly invite all stakeholders, particularly teachers and students, to contribute to the Southeast Asian Journal on STEM Education. Your contributions to research, practices, and reflections do more than enrich our collective knowledge; they play an active role in sculpting

a future where STEM education is universally accessible, inclusive, and profoundly influential. This journal stands as a testament to the strength of collaborative effort, reinforcing the notion that every contribution, regardless of its size, is instrumental in unlocking the vast potential of STEM.

The Southeast Asian Journal on STEM Education serves as a beacon, illuminating the path toward a more enlightened future via the progression of STEM education. It summons all who are dedicated to nurturing scientific thought and leveraging scientific inquiry to elevate the quality of teaching and learning. United in purpose, let us venture forth on this path of exploration and innovation, laying the foundations for future generations to build upon.

Best regards,

SAJSE Editors

POLICY

SECTION

OUT-OF-SCHOOL CHILDREN, YOUTH AND ADULT (OSCYA) LEARNERS STEM PERFORMANCE QUANTIFIER AND SENTIMENT: INSIGHT FROM ACCREDITATION AND EQUIVALENCY (A&E) TEST

Dionito F Mangao Jr¹

*Alternative Learning System, Schools Division of Cavite Province, Region IV-A,
Department of Education, Philippines
<dionitojr.mangao@deped.gov.ph>*

ABSTRACT

This research investigates the STEM literacy performance and sentiments of Out-of-School Children, Youth and Adults (OSCYA) based on the results of the 2024 Accreditation and Equivalency (A&E) Test. Specifically, the study examines learners' performance across six learning strands and explores the relationship between learner sentiment and achievement in core STEM areas within the context of non-formal education, particularly Mathematics, Science and Technology. A mixed descriptive research design was employed, integrating mean percentage scores with inferred sentiment categories—positive, neutral and negative—based on learners' attitudes towards STEM. The findings revealed that Science and Digital Literacy (Technology) demonstrated moderate performance, with mean scores ranging from 61.19% to 68.73%. However, Mathematics consistently recorded the lowest performance across both Elementary and Junior High School levels, with mean scores of 44.97% and 39.26%, highlighting substantial gaps in numeracy and problem-solving skills. In contrast, non-STEM learning strands, particularly Life and Career Skills and Understanding Self and Society, yielded the highest performance, indicating that OSCYA learners demonstrate stronger competencies in applied and real-world domains. Overall, the results underscore the need for differentiated and learner-centred interventions that address both the academic and affective needs of OSCYA learners. Strengthening STEM literacy requires the provision of differentiated instruction, particularly in Mathematics, alongside contextualised and practical learning experiences that connect STEM concepts to learners' everyday lives, work and community contexts. Strategies aimed at enhancing learner engagement, building confidence and fostering positive experiences in STEM academic performance. In particular, targeted numeracy and problem-solving programmes, supportive teacher feedback, collaborative learning activities and authentic applications of Science, Technology and Mathematics should be prioritised. The study concludes that integrating affective factors into educational planning and instructional strategies is essential for strengthening STEM literacy among marginalised out-of-school children, youth and adults and for promoting more equitable learning opportunities within non-formal education.

Keywords: Out-of-School, Non-Formal Education, Alternative Learning System (ALS), Sentiment, STEM Performance, Accreditation and Equivalency (A&E)

1. Introduction

Education is recognised globally as a fundamental human right and a critical factor for economic growth, social inclusion, and sustainable development. However, Access to quality education remains a fundamental tool for social mobility and national development, yet many countries continue to face challenges in addressing educational exclusion among out-of-school children, youth, and adults (OSCYA). In Southeast Asia OSCYA remains one of the most pressing educational concerns despite significant improvements in school access over

recent decades. In addition, millions of children and adolescents across Southeast Asia continue to face barriers to education due to poverty, geographical isolation, social inequality, disability, migration, and inadequate educational infrastructure. According to the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2024), approximately 251 million children and youth are still out of school, emphasising the urgent need for targeted educational interventions across the world.

Furthermore, the educational exclusion experienced of out-of-school children varies significantly among Southeast Asian countries, including the Philippines, Indonesia, Cambodia, Myanmar, and Laos. While governments introduced reforms to increase enrolment and improve educational access, marginalised populations remain disproportionately affected. Children from low-income families, indigenous communities, and geographically isolated areas are more likely to experience interrupted schooling or non-enrolment (Inui et.al., 2025). This educational disparity highlights persistent structural inequalities that continue to hinder progress toward universal access to quality education.

Moreover, several interconnected factors contribute to the high number of out-of-school children and youth in Southeast Asia. Poverty remains the most significant barrier, as many children are forced to prioritise income-generating activities or household responsibilities over schooling. Limited access to schools in rural and conflict-affected areas, inadequate educational resources, and insufficient policy implementation further aggravate the situation (UNESCO, 2017). Additionally, the COVID-19 pandemic intensified existing educational inequalities by disrupting learning systems and increasing dropout rates, particularly among vulnerable populations (UNESCO, 2023). These challenges underscore the complexity of addressing educational exclusion in the region.

In the Philippines context, the Alternative Learning System (ALS) serves as a crucial intervention in providing second-chance education, but not a second class to learners who have not completed formal basic education. Through the Accreditation and Equivalency (A&E) programme by the Department of Education Alternative Learning System, OSCYA learners are given opportunities to take A&E test equivalent to formal school graduates and reintegrate into higher education, technical-vocational pathways, or employment opportunities (DepEd, 2022). Despite the significant effort of ALS implementers, concerns remain regarding learners' academic performance, particularly in Science, Technology, Engineering, and Mathematics (STEM), which are increasingly recognised as critical fields for global competitiveness and innovation-driven economies. STEM education has become an essential component of 21st century learning because it equips individuals with analytical, problem-solving, and technological skills necessary for participation in modern society. According to Darling-Hammond et. al. (2020) strong STEM competencies are directly linked to workforce readiness, economic growth, and sustainable development. However, marginalised learners, including those in ALS alternative education pathways, often face barriers such as limited access to resources, insufficient instructional support, learners' confidence, motivations and interest and gaps in prior foundational knowledge. These challenges can significantly affect their academic performance, particularly in Accreditation and Equivalency (A&E) examination, where mastery of STEM-related competencies is increasingly emphasised and a major concern.

Although the present study is situated within the Philippines Alternative Learning System, its policy significance extends to the broader Southeast Asian context. Many countries in the region continue to develop alternative, flexible, or non-formal pathways for learners whose participation in formal education has been interrupted. These systems face a common challenge: providing equitable access to foundational competencies while also preparing learners for further education, employment, digital participation, and rapidly changing labour-market demands. From this perspective, evidence from the Philippine ALS experience can provide a useful case for considering how assessment data may inform STEM policy for marginalised learners across Southeast Asia. The findings should not be interpreted as

representative of the region as a whole, but rather as a country-level case that highlights policy issues that may warrant comparative attention across Southeast Asian education systems.

In relation to ALS 2024 A&E results provide an important opportunity to examine the OSCYA ALS learners' STEM performance and identify patterns that may inform non-formal educational interventions. Quantifying learners' STEM performance offers measurable evidence of strengths and areas requiring critical enhancement and improvement and allowing educators and policymakers to evaluate the effectiveness of ALS instructional teaching strategies. Previous studies have shown that assessment data can serve as a valuable tool for identifying achievement gaps and designing responsive educational programmes (Darling-Hammond et al., 2020). For ALS, analysing learners' STEM performance is a major concern and critical because it can reveal whether current learning modalities and strategies adequately prepare OSCYA learners to meet academic and professional demands in STEM-related fields.

Finally, beyond academic performance, learners' sentiments and perceptions toward STEM learning play a significant role in shaping educational outcomes. Research suggests that positive learner sentiment, including confidence, motivation, and perceived relevance of STEM subjects, contributes to higher academic achievement and persistence (Eccles & Wigfield, 2020). Conversely, negative perceptions such as anxiety, low self-efficacy, and lack of interest may hinder performance. For OSCYA learners, who often juggle socioeconomic challenges, work responsibilities, and interrupted educational experiences, understanding their sentiments toward STEM can provide deeper insight into the factors influencing their examination performance. Examining both quantitative performance metrics and qualitative learner sentiments allows for a more comprehensive understanding of their educational realities. In addition, this study seeks to investigate the relationship between STEM performance and learner sentiment among ALS participants based on the 2024 A&E examination results. By integrating performance quantification with sentiment analysis, the research aims to generate evidence-based insights that can support curriculum enhancement, targeted instructional interventions, and policy improvements within the ALS framework and policies. Ultimately, the findings of this study may contribute to strengthening non-formal education efforts and ensuring that OSCYA learners are better equipped to succeed in an increasingly STEM-driven world.

2. Objectives of the Study

The objective of the study aims to answer the following question:

- a) Research Question 1: What is the level of STEM performance of OSCYA learners in the 2024 Accreditation and Equivalency (A&E) test in terms of Science (Critical Thinking), Math (Mathematical and Problem-Solving Skills), and Technology (Digital Literacy) across Elementary and Junior High School levels?
- b) Research Question 2: Is there a significant difference in STEM performance between Elementary and Junior High School Level OSCYA learners A&E results?
- c) Research Question 3: How does learners' sentiment (positive, neutral, and negative) influence their performance in STEM-related learning strands, particularly in Mathematics, Science, and Technology?

3. Methodology

3.1 Research Design

This study employed a descriptive mixed method approach to examine the STEM performance of Out-of-School Children, Youth, and Adult (OSCYA) learners based on the 2024 Accreditation and Equivalency (A&E) examination results, alongside sentiment analysis to assess learners' attitudes toward STEM-related learning strands. The descriptive approach was utilised to systematically present and interpret numerical data on learners' academic performance in Science, Mathematics, and Digital Literacy (Technology), while the quantitative component enabled objective measurement and statistical analysis of learner achievement. In addition, the study incorporated a comparative and correlational approach to determine performance variations across educational levels and to analyse the relationship between learner sentiment and STEM achievement. The qualitative responses were categorised into positive, neutral, and negative sentiment classifications based on learners' attitudes, motivation, confidence levels, and perceptions regarding STEM subjects.

3.2 Participants of the Study

The participants of the study consisted of Out-of-School Children, Youth, and Adult (OSCYA) learners who took the 2024 Accreditation and Equivalency (A&E) examination under the Alternative Learning System (ALS) in the Department of Education, Division of Cavite Province, Philippines. The participants were categorised into two groups: Elementary Level and Junior High School Level learners. The study employed total population sampling, wherein all available A&E examination records relevant to STEM-related learning strands were included in the analysis. The participants represented diverse backgrounds, including individuals who had interrupted formal schooling due to socioeconomic difficulties, family responsibilities, geographical barriers, and other personal circumstances. For the sentiment analysis component, learner responses collected through available feedback mechanisms were classified according to their expressed perceptions toward STEM learning.

3.3 Data Gathering Method

The study utilised secondary data analysis as the primary data gathering method. Official records of the 2024 A&E examination results were obtained from the Alternative Learning System database, focusing specifically on STEM-related learning strands such as Scientific and Critical Thinking Skills, Mathematical and Problem-Solving Skills, and Digital Literacy (Technology). The collected data included passed, conditional, and failure rates for both Elementary and Junior High School learners. For sentiment analysis, learner feedback, reflection responses were gathered and reviewed. The qualitative responses were categorised into positive, neutral, and negative sentiment classifications based on learners' attitudes, motivation, confidence levels, and perceptions regarding STEM subjects. Prior to analysis, all collected data were carefully organised, validated, and encoded for statistical treatment.

3.4 Data Analysis

The data were analysed using descriptive and inferential statistical techniques. Descriptive statistics such as mean percentage scores, frequency counts, and percentages were used to summarise learner performance across the identified STEM learning strands. Comparative analysis was conducted to determine differences in performance between Elementary and Junior High School learners. Performance levels were interpreted using established percentage scales to classify achievement as high, moderate, or low. For qualitative insight, sentiment analysis, learner responses were grouped into positive, neutral, and negative categories, and corresponding mean scores in Mathematics, Science, and Digital Literacy

(Technology) were compared to identify performance trends. Furthermore, correlational analysis was employed to determine the relationship between learners' sentiment categories and their STEM performance outcomes.

4. Result and Discussion

Table 1 shows the Elementary Level (EL) A&E Result 2024 (Mean Percentage by Learning Strand) reveal varying levels of academic performance across the six learning strands among the 662 examinees, indicating both strengths and areas that require targeted intervention within the elementary level assessment. Among all strands, Life and Career Skills recorded the highest passing percentage at 87.16%, suggesting that learners demonstrate strong practical competencies and applied life skills. This reflects the effectiveness of instruction in developing real-world readiness and adaptive capabilities among learners. However, the relatively high failure rate of 56.77% in the same strand indicates inconsistencies in learner mastery, suggesting that while some learners excel significantly, others struggle to meet the required competency standards.

Table 1

Elementary Level A&E Result 2024 (Mean Percentage by Learning Strand)

Learning Strand	Passed (%)	Conditional (%)	Failed (%)	Interpretation
1-English	66.14	66.14	32.39	Average performance with moderate failure rate
1-Filipino	80.92	80.92	52.91	High performance but notably high failure proportion
2-Scientific and Critical Thinking (<i>Science</i>)	66.47	66.47	29.56	Moderate STEM competency
3-Mathematical and Problem Solving (<i>Math</i>)	44.97	44.97	27.21	Low performance in core STEM skill
4-Life and Career Skills	87.16	87.16	56.77	Highest performance but inconsistent outcomes
5-Understanding Self and Society	81.76	81.76	49.00	Strong social understanding
6-Digital Literacy (<i>Technology</i>)	68.73	68.73	29.03	Moderate digital/STEM-related skills

Understanding self and Society obtained a high passing rate of 81.76%, indicating that learners possess strong awareness of social relationships, personal development, and civic understanding. This result suggests that the curriculum effectively fosters socio-emotional and societal competencies. Filipinos also showed strong performance, with a passing rate of 80.92%, reflecting learners' relative proficiency in the national language and comprehension skills. However, the notably high failure rate of 52.91% raises concerns regarding uneven mastery levels, which may stem from differences in language exposure, instructional delivery, or literacy foundations among learners. In contrast, English posted a moderate passing percentage of 66.14%, with a failure rate of 32.39%, indicating average learner performance. This suggests that while a considerable number of examinees demonstrate functional English proficiency, there remains a significant portion who may require additional support in reading

comprehension, vocabulary development, and communication skills. Likewise, Scientific and Critical Thinking (*Science*) recorded a passing rate of 66.47%, indicating moderate competency in scientific reasoning and analytical thinking. This performance suggests that learners possess foundational understanding of scientific concepts but may need enhanced opportunities for inquiry-based learning and critical analysis exercises. The results further reveal challenges in Digital Literacy, which achieved a passing rate of 68.73%. While this reflects moderate competency in digital-related skills, the findings indicate that a substantial number of learners still face difficulties in navigating digital tools and technology-based tasks. Given the increasing importance of digital competence in modern education and employment, strengthening this strand should be considered a priority.

Most notably, Mathematical and Problem-Solving Skills recorded the lowest passing percentage at 44.97%, making it the weakest-performing strand among all assessed areas. This indicates significant difficulties in numeracy, logical reasoning, and problem-solving competencies among learners. The low performance in mathematics may be attributed to gaps in foundational concepts, limited practice opportunities, or insufficient learner confidence in mathematical tasks. Overall, in Elementary level, the findings suggest that learners demonstrate stronger performance in practical, social, and language-related strands, while showing weaker outcomes in mathematics and moderate competency in science and digital literacy. These results highlight the need for targeted instructional interventions, particularly in strengthening STEM-related competencies, to ensure balanced learner development across all learning strands.

Table 2

Junior High School Level (JHSL) A&E Result 2024 (Mean Percentage by Learning Strand)

Learning Strand	Passed (%)	Conditional (%)	Failed (%)	Interpretation
1-English	74.63	60.25	38.87	Improved performance vs. elementary
1-Filipino	74.48	60.86	41.90	Slight decline in mastery consistency
2-Scientific and Critical Thinking (<i>Science</i>)	61.19	43.54	29.26	Moderate STEM proficiency
3-Mathematical and Problem Solving (<i>Math</i>)	39.26	31.04	25.39	Persistently lowest across strands
4-Life and Career Skills	87.54	79.69	56.98	Consistently strongest area
5-Understanding Self and Society	85.73	75.02	50.08	High performance maintained
6-Digital Literacy (<i>Technology</i>)	67.61	48.05	28.37	Stable but not advanced STEM related skill

Furthermore, table 2 shows the Junior High School Level (JHSL) A&E Result 2024 reveals varying levels of student performance across the different learning strands, highlighting both strengths and areas that require intervention. Among all strands, Life and Career Skills recorded the highest percentage of passed learners at 87.54%, with a relatively strong conditional rate of 79.69%, indicating that students demonstrate solid competence in practical and adaptive life skills. Similarly, Understanding Self and Society performed strongly, with 85.73% passing, suggesting that learners possess a good understanding of personal development and social awareness. These results reflect the students' strong grasp of applied and socially relevant competencies. In language-related strands, both English and Filipino posted satisfactory passing rates of 74.63% and 74.48%, respectively. English showed a

slightly stronger overall performance and an indication of improved mastery compared to the elementary level, while Filipino reflected a slight decline in consistency, as shown by the higher failed percentage (41.90%). This suggests that while students have generally developed adequate communication skills, reinforcement in Filipino language proficiency may be necessary to improve mastery and retention. The performance in Scientific and Critical Thinking (Science) yielded a moderate passing rate of 61.19%, which suggests that learners have a foundational understanding of scientific concepts but may still struggle with higher-order analytical and critical thinking tasks. Likewise, Digital Literacy recorded a pass rate of 67.61%, indicating that students possess basic technological competencies but have yet to reach advanced STEM-related skill levels. The findings point out the need for more engaging, inquiry-based, and technology-integrated learning strategies to further strengthen students' scientific reasoning and digital competence. The most significant concern is observed in Mathematical and Problem Solving, which recorded the lowest passing rate at only 39.26%. This strand also has relatively low conditional performance (31.04%), confirming that mathematics remains the weakest area among JHS learners. The result suggests persistent difficulties in numerical reasoning, analytical thinking, and problem-solving application. This calls for immediate instructional intervention, such as remediation programmes, differentiated teaching strategies, and intensified practice exercises to address learning gaps and improve mathematical proficiency. The JHS results indicate that learners perform better in practical life skills, social understanding, and language competencies, while demonstrating weaker performance in mathematics and moderate achievement in STEM-related areas. These findings emphasise the importance of targeted academic support, particularly in mathematics and science, while sustaining the strengths observed in life skills and social development strands to ensure balanced learner achievement across all competencies.

In comparison, both Elementary and JHS A&E results revealed notable trends in learner performance across different learning strands, showing areas of improvement as learners' progress to higher levels, as well as persistent challenges that remain evident. Overall, the results suggest that while JHS learners generally demonstrate improved mastery in several strands compared to elementary learners, some foundational academic difficulties—particularly in mathematics—continue affect learners' performance across both levels. In the language strands, JHS learners showed improvement in English, with the passing rate increasing from 66.14% at the elementary level to 74.63% at the junior high school level. This indicates stronger language comprehension and communication skills as learners advance. However, Filipinos showed a slight decline, decreasing from 80.92% in elementary to 74.48% in junior high school. This suggests that while elementary learners demonstrate stronger Filipino proficiency, maintaining mastery and consistency becomes more challenging at the higher level, possibly due to increased complexity of language tasks and comprehension requirements. For Scientific and Critical Thinking (Science), there was a slight decline from 66.47% at the elementary level to 61.19% at the junior high school level. This indicates that although learners possess moderate STEM competency in both levels, the more advanced scientific concepts and analytical demands in junior high school may pose difficulties. Similarly, Digital Literacy dropped marginally from 68.73% to 67.61%, suggesting that while basic digital competencies are maintained, learners may struggle to transition toward more advanced technological and STEM related applications. The most consistent area of concern across both levels is Mathematical and Problem Solving. The passing rate declined from 44.97% at the elementary level to 39.26% in junior high school, making it the lowest-performing strand in both categories. This persistent underperformance indicates significant gaps in mathematical reasoning and problem-solving skills, which become more pronounced as mathematical concepts become more abstract and complex in higher levels. The results strongly highlight the need for targeted interventions, enhanced instructional strategies, and sustained remediation programmes in mathematics across both educational stages.

On the other hand, Life and Career Skills remained the strongest-performing strand in both levels, with a slight increase from 87.16% in elementary to 87.54% in junior high school. This

consistency reflects learners' strong practical and adaptive competencies, suggesting that experiential and contextual learning approaches are effective in this area. Likewise, Understanding Self and Society improved from 81.76% to 85.73%, indicating that learners continue to develop social awareness, self-understanding, and civic competence as they progress academically. In terms of learner population, the junior high school level had significantly more takers (2,793) compared to the elementary level (662), which may contribute to broader performance variation. Despite this, the overall trend shows that learners tend to strengthen in applied life skills, social understanding, and English proficiency as they advance, while challenges persist in mathematics and certain STEM-related competencies.

Overall, the comparison reveals that progression from elementary to junior high school generally results in improved performance in practical, social, and English-related strands, but also exposes persistent academic weaknesses in mathematics and moderate declines in science and digital literacy. These findings emphasise the need for curriculum enhancement, stronger foundational STEM instruction, and differentiated support systems to ensure that learners can achieve balanced academic growth across all learning strands.

Table 3

Summary of STEM-Related Performance (OSCYA Focus)

STEM Component	Elementary Mean (%)	JHS Mean (%)	Overall Interpretation
LS2-Scientific and Critical Thinking (<i>Science</i>)	66.47	61.19	Moderate but declining at higher level
LS3-Mathematical and Problem Solving (<i>Math</i>)	44.97	39.26	Consistently weakest STEM area
LS6-Digital Literacy (<i>Technology</i>)	68.73	67.61	Stable but not high performing

The Summary of STEM-Related Performance (OSCYA Focus) in table 3 highlights important trends in learners' achievement across key STEM components—Science (Critical Thinking), Mathematics, and Digital Literacy—from the elementary level to the junior high school level. The results reveal that while learners demonstrate moderate competence in some STEM-related areas, there is a noticeable decline in performance as they transition to higher educational levels, suggesting increasing difficulty in mastering more complex STEM concepts and skills. In Science (Critical Thinking), the mean percentage decreased from 66.47% at the elementary level to 61.19% at the junior high school level. This decline indicates that although learners possess a fair understanding of basic scientific concepts and critical thinking processes, they may struggle to adapt to the more advanced analytical, investigative, and problem-solving demands required at the junior high school level. The results suggest a need for more inquiry-based learning approaches, laboratory-based instruction, and activities that strengthen higher order thinking skills to help learners better engage with increasingly complex scientific content. The most critical result is observed in Mathematics, which recorded the lowest performance among all STEM components at both levels, declining from 44.97% in elementary to 39.26% in junior high school. This consistent underperformance confirms that mathematics remains the weakest STEM area for learners. The decline suggests that foundational mathematical gaps established at the elementary level may continue to affect learners as they encounter more abstract and advanced mathematical concepts in junior high school. This calls for urgent intervention through strengthened numeracy programmes, targeted remediation, differentiated instruction, and more learner-centred teaching strategies that focus on conceptual understanding and practical application of mathematical skills.

Meanwhile, Digital Literacy showed relatively stable performance, with only a slight decrease from 68.73% at the elementary level to 67.61% at the junior high school level. This indicates that learners can maintain basic digital competencies across levels, reflecting a consistent exposure to technology-related skills. However, the results also suggest that learners are not progressing toward higher-level digital proficiency, such as advanced technological application, computational thinking, and digital problem-solving. This stability, while positive, points to the need for curriculum enhancement that moves beyond basic digital skills toward more advanced and innovative STEM integration.

Overall, the findings suggest that STEM-related performance among learners is generally moderate but shows signs of decline at higher levels, particularly in science and mathematics. Mathematics remains the most critical area for improvement, while science requires stronger instructional support to sustain critical thinking development. Digital literacy, though stable, needs further advancement to meet evolving technological demands. These results emphasise the importance of strengthening STEM education through enhanced instructional strategies, improved learning resources, and focused intervention programmes to ensure that learners develop the competencies necessary for future academic and career success in STEM-related fields.

The persistent underperformance in Mathematical and Problem-Solving Skills warrants particular attention as a policy priority rather than being treated solely as an instructional concern. The decline from 44.97% at the elementary level to 39.26% at the junior high school level suggests that foundational gaps may not be adequately identified and addressed as learners progress through the ALS pathway. This pattern points to the need for a system-level approach to numeracy that includes early diagnostic assessment, structured remediation, differentiated learning materials, continuous monitoring of learner progress, and professional development for teachers supporting learners with diverse mathematical backgrounds. Strengthening mathematics in ALS may therefore require coordinated action across curriculum, assessment, teacher support, learning resources, and community-based learning provision rather than reliance on additional practice alone.

Table 4

Relationship between Sentiment and STEM Literacy Achievement (Focus on Mathematics)

Sentiment Category	Math (%)	Science (%)	Digital Literacy (%)	Interpretation
Positive	52.30	70.15	72.48	Learners with positive attitudes perform better across STEM, especially in problem-solving tasks
Neutral	44.10	63.83	68.02	Moderate performance; reflects uncertainty or limited engagement in STEM
Negative	38.75	58.60	64.11	Lowest performance, especially in Math, indicating strong association with poor outcomes

The pattern observed across sentiment categories provides an important affective dimension to the interpretation of STEM achievement. Learners classified as having positive sentiments demonstrated higher mean performance across Mathematics, Science, and Digital Literacy than those in the neutral and negative categories. However, because the present study is

based on descriptive and correlational evidence, these findings should be interpreted as an association rather than evidence that sentiment directly causes higher or lower achievement. The pattern nevertheless suggests that learner confidence, motivation, perceived relevance, and engagement warrant consideration in the design of STEM support within non-formal education. Further research using longitudinal or intervention-based designs would be valuable in determining whether targeted approaches to strengthening STEM confidence and engagement can contribute to improved achievement among OSCYA learners.

Finally, the results in Table 4 show a clear relationship between learners' sentiment toward mathematics and their STEM literacy achievement. Students with positive sentiments obtained the highest mean scores across all areas, with 52.30% in Mathematics, 70.15% in Science, and 72.48% in Digital Literacy. These findings suggest that learners who display positive attitudes and confidence toward mathematics tend to perform better not only in Math but also in other STEM-related domains. A positive mindset may encourage greater engagement, persistence, and willingness to solve complex problems, which contributes to higher academic achievement. In contrast, learners with neutral sentiments demonstrated moderate performance, obtaining mean scores of 44.10% in Mathematics, 63.83% in Science, and 68.02% in Digital Literacy. This indicates that students who are uncertain or less emotionally engaged with mathematics still perform reasonably but may lack the motivation or confidence needed to excel fully in STEM learning. Their results suggest a moderate level of understanding and participation compared to students with positive attitudes. Meanwhile, learners with negative sentiments recorded the lowest performance in all areas, particularly in Mathematics, where the mean score dropped to 38.75%. Their Science and Digital Literacy scores were also lower at 58.60% and 64.11%, respectively. The findings imply that negative attitudes toward mathematics are strongly associated with weaker STEM outcomes. Learners who experience fear, frustration, or lack of interest in mathematics may struggle with problem-solving tasks and show lower engagement in STEM-related activities. Overall, the data reveal a consistent trend in which more positive sentiments correspond to higher STEM literacy achievement. This highlights the importance of fostering positive learning experiences, motivation, and confidence in mathematics to improve students' overall STEM performance.

The relatively stable performance in Digital Literacy also presents an important policy consideration. Maintaining basic digital competence is valuable, but the findings suggest that ALS provision may need to move beyond functional technology use toward higher-order digital competencies, including information evaluation, digital problem-solving, computational thinking, and responsible use of technology. For marginalised learners, this progression is particularly important because digital competence increasingly influences access to education, employment, public services, and participation in society. Policy responses should therefore consider not only access to digital devices and connectivity but also the quality, relevance, and progression of digital learning opportunities available through Alternative Learning System (ALS) program and Community Learning Centers (CLCs).

5. Policy Implications for Inclusive STEM Education in Alternative Learning Systems

The findings of this study have implications beyond the assessment of individual learner performance. They highlight structural and pedagogical issues that should be considered in strengthening STEM provision for learners participating in alternative and non-formal education. The consistently low performance in Mathematics, the moderate achievement in Science, the limited progression in Digital Literacy, and the association between learner sentiment and STEM achievement collectively suggest that a comprehensive policy response is required. Such a response should recognise that OSCYA learners have diverse educational histories, learning needs, socioeconomic circumstances, and levels of prior knowledge.

Consequently, policies for STEM education in alternative learning pathways should prioritise flexibility, differentiation, foundational learning, learner engagement, and equitable access to relevant STEM learning opportunities.

5.1 Strengthening Foundational STEM Learning

First, foundational STEM learning—particularly numeracy and mathematical problem-solving—should be established as a priority within ALS improvement efforts. The consistently lowest mathematics performance across both educational levels indicates that learning gaps may accumulate as learners progress. A policy response could include systematic diagnostic assessment, targeted numeracy programmes, differentiated learning pathways, structured remediation, and regular monitoring of learner progress. Such measures would allow ALS providers to identify learning gaps earlier and provide support according to individual learner needs rather than relying primarily on grade-level or examination-based expectations.

5.2 Teacher Professional Development

Second, strengthening teacher capacity should form an integral part of STEM policy for non-formal education. ALS teachers work with learners who may differ substantially in age, prior schooling, confidence, motivation, and learning pace. Professional development should therefore equip teachers with strategies for differentiated STEM instruction, contextualised problem-solving, formative assessment, inquiry-based learning, and approaches that build learner confidence. Teacher support should also include access to adaptable STEM teaching resources that can be used in flexible and community-based learning environments.

5.3 Data-Informed Policy and Monitoring

Third, assessment data should be used more systematically to inform policy planning and resource allocation. The A&E results demonstrate how assessment evidence can identify persistent weaknesses and differences across STEM learning strands and educational levels. Establishing mechanisms for regularly analysing STEM performance data could support the identification of priority learning gaps, guide the allocation of learning resources, inform teacher development, and monitor the effectiveness of targeted interventions. Importantly, assessment data should be used not only to measure outcomes but also to support timely and responsive educational decision-making.

5.4 Learner-Centred and Inclusive STEM Policy

Finally, STEM policy for OSCYA learners should incorporate both cognitive and affective dimensions of learning. The association between positive sentiment and higher STEM performance suggests that confidence, motivation, engagement, and perceived relevance should be considered alongside academic competencies. Policies and programmes should therefore promote supportive learning environments, contextualised STEM activities, collaborative learning, positive feedback, and opportunities for learners to experience meaningful success. Such approaches can help position STEM not simply as a set of academic subjects but as a pathway through which marginalised learners can develop confidence, practical competencies, and opportunities for further education and employment.

From a Southeast Asian perspective, the Philippines experience highlights the importance of strengthening STEM within alternative and second-chance education pathways as part of broader efforts to advance educational equity and lifelong learning. While the findings of this study are specific to OSCYA learners in Cavite Province and should not be generalised directly to other national contexts, they raise policy questions relevant across the region: how alternative education systems identify and respond to foundational STEM gaps; how teachers

are prepared to support diverse learners; how digital competencies are developed beyond basic skills; and how assessment evidence can be translated into targeted policy action. Comparative research across Southeast Asian countries could build on these questions and contribute to a stronger regional evidence base for inclusive STEM policy.

6. Conclusion

Overall, this study demonstrates that STEM achievement among OSCYA learners should be understood not only as an issue of individual academic performance but also as a matter of educational equity and system responsiveness. The persistent weakness in Mathematics, moderate performance in Science, limited progression in Digital Literacy, and observed association between learner sentiment and STEM achievement point to the need for a more differentiated and learner-centred approach to STEM provision within alternative education. For policymakers and programme implementers, the findings support greater emphasis on foundational numeracy, data-informed intervention, teacher professional development, progressive digital literacy, and supportive learning environments. Although the evidence is drawn from one Philippine context, the study provides a useful case for wider Southeast Asian dialogue on how non-formal education systems can ensure that marginalised learners have meaningful opportunities to develop the STEM competencies required for further learning, employment, and participation in an increasingly technology-oriented society.

The findings therefore recommend the strong implementation of learner-centred and supportive interventions, particularly in Mathematics, such as differentiated instruction, hands-on problem-solving activities, collaborative learning, real-life applications, and regular formative assessment in ALS learning sessions. In addition, ALS teachers should also provide positive feedback, strong encouragement, and opportunities for ALS learners to experience small but meaningful successes in order to improve confidence and reduce anxiety toward STEM subjects' difficulty. For Science, inquiry-based experiments and contextualised activities may help sustain interest and improve conceptual understanding, while Digital Literacy should be enhanced through authentic technology-based projects that develop not only basic computer skills but also critical thinking, information evaluation, digital problem-solving, and responsible technology use.

Finally, the results demonstrate that STEM achievement is closely connected to both academic competency and learner sentiment. The decline from elementary to JHS performance in Science and Mathematics suggests the need for stronger continuity in curriculum implementation, early identification of learning gaps, and targeted ALS remediation programmes particularly in Mathematics should be prioritised through structured intervention programmes, teacher professional development, peer tutoring, and sustained practice in higher-order thinking and problem-solving. At the same time, the positive performance associated with favourable sentiment highlights the value of creating an encouraging, inclusive, and motivating learning environment. Community Learning Centers (CLCs) should therefore adopt a comprehensive approach that combines academic support with strategies for improving learner attitudes, confidence, and engagement. By addressing both the cognitive and affective dimensions of learning, educational stakeholders, local government unit participation can also help to improve ALS program implementation on the enhancement of learners STEM performance for accreditation and equivalency test, strengthen mathematical, scientific and critical thinking reasoning, maintain digital literacy gains, and ultimately develop learners who are more capable, motivated, and prepared to participate successfully in an increasingly STEM-oriented society and academic success of Out-of-School Children, Youth and Adult marginalised learners.

References

- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, 24(2), 97–140. <https://doi.org/10.1080/10888691.2018.1537791>
- Department of Education. (2022). *Implementing rules and regulations of the Alternative Learning System Act (Republic Act No. 11510)*. <https://elibrary.judiciary.gov.ph/thebookshelf/showdocs/2/92765>
- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, Article 101859. <https://doi.org/10.1016/j.cedpsych.2020.101859>
- Inui, M., Ogisu, T., Kamogawa, A., & Nakaya, A. (2025). Out-of-school children in Southeast Asia: A case study of Cambodia, Laos, Indonesia, and Malaysia. *Cogent Education*, 12(1), Article 2481004. <https://doi.org/10.1080/2331186X.2025.2481004>
- UNESCO. (2017). *Situation analysis of out-of-school children in nine Southeast Asian countries*. UNESCO Bangkok. <https://unesdoc.unesco.org/ark:/48223/pf0000252749/PDF/252749eng.pdf.multi>
- UNESCO. (2023). *Global education monitoring report 2023: Technology in education—A tool on whose terms?* UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000385723>
- UNESCO. (2024). *251 million children still out of school worldwide, UNESCO reports*. <https://www.unesco.org/en/articles/251m-children-and-youth-still-out-school-despite-decades-progress-unesco-report>

TEACHER

SECTION

GAUTHMATH AS AI-ASSISTED SCAFFOLDING IN SECONDARY MATHEMATICS: LEARNER ACHIEVEMENT AND ACCEPTABILITY IN COMBINATORICS

Christian Genesis R Guinto*
Maria Magdalena S Cabangca

*Department of Mathematics, Cotta National High School, Schools Division of Lucena City,
Philippines*
<christiangenesis.guinto@deped.gov.ph>

ABSTRACT

This action research examined the effects of the Gauthmath application, an AI-assisted mathematics platform, on the achievement and acceptability of Grade 10 learners ($n = 30$) in permutations and combinations at a national high school in the Philippines (SY 2024–2025). The teacher-researcher employed a one-group pre-test–post-test design over a five-week unit (25 sessions, 18.75 hours), aligned with the prescribed Most Essential Learning Competency (MELC) schedule: permutations (Weeks 1–2), combinations (Weeks 3–4), and integration (Week 5). Grounded in Vygotsky's (1978) Zone of Proximal Development (ZPD), Sweller's (1988) and Sweller, van Merriënboer, and Paas's (2019) Cognitive Load Theory (CLT), and Ryan and Deci's (2000) Self-Determination Theory (SDT), the study explored how AI-assisted step-by-step scaffolding supports combinatorics learning as a Technology–Mathematics STEM integration within the Community and School sub-theme. Each daily session followed a three-phase protocol: an independent problem attempt, an application-assisted review using Gauthmath's annotated solutions, and a structured reflection. Results showed strongly positive learner acceptability across usability (Weighted Average Mean [WAM] = 4.58), relevance (WAM = 4.63), and engagement (WAM = 4.69), all interpreted as Strongly Agree. A statistically significant performance gain was observed from pre-test (Mean Percentage Score [MPS] = 27.44%) to post-test (MPS = 81.56%), $t(29) = -22.4$, $p < .001$, $d = 4.09$. Five of six MELC competency clusters reached mastery at post-test; circular and distinguishable permutations (73.3%) remained approaching the threshold. After adjusting for regression-to-mean artefact (approximately 24% of the gain), a large residual gain of approximately 41 percentage points persists. The findings support the daily integration of Gauthmath as a ZPD-aligned supplement during the combinatorics MELC unit in Philippine secondary mathematics classrooms.

Keywords: Gauthmath, Combinatorics, AI-Assisted Scaffolding, MELC, Action Research

1. Introduction

Filipino secondary learners continue to record low performance in mathematics, a gap that has persisted rather than improved across successive assessment cycles. In its first participation in the Programme for International Student Assessment (PISA) in 2018, the Philippines scored 353 in mathematics, among the lowest internationally (Department of Education, 2019); by the 2022 cycle, the mean mathematics score had moved only marginally to 355, still among the five lowest-performing nations internationally (OECD, 2023). At the competency level, curriculum-mapping analysis by Gratela et al. (2023) identified permutations and combinations as among the few Grade 10 mathematics competencies with no introductory treatment in earlier grade levels, unlike most other strands that are reinforced across the K-to-10 sequence, a structural gap that plausibly compounds the topic's intrinsic

difficulty. Post-pandemic disruptions have further widened these gaps, disproportionately affecting learners in under-resourced classrooms (Ando et al., 2022; Mishra et al., 2020).

The Gauthmath application (Gauthtech Pte. Ltd., Singapore) is an AI-assisted mathematics platform that accepts problem inputs via photograph capture or text entry and returns step-by-step annotated solutions through an Optical Character Recognition (OCR) and Symbolic Reasoning Engine. For combinatorics, it identifies the problem type, selects the governing formula, and presents the computation in labelled steps with natural-language rationales. This operationalises the scaffolded support that Vygotsky's (1978) Zone of Proximal Development (ZPD) identifies as the driver of learning beyond independent capacity: the application delivers just-in-time feedback calibrated to the learner's specific submitted problem, enabling progressive internalisation of combinatorial schemas across daily sessions (Belland et al., 2017).

This scaffolding mechanism is further grounded in Cognitive Load Theory (CLT), drawing on Sweller (1988) and Sweller, van Merriënboer, and Paas (2019). Combinatorics carries exceptionally high intrinsic load; Gauthmath's consistently structured, step-annotated interface reduces extraneous load, freeing working memory for germane schema formation. Ryan and Deci's (2000) Self-Determination Theory (SDT) provides the motivational account: the application supports learner autonomy (self-paced problem selection), competence (immediate accurate feedback), and relatedness (peer solution discussion), sustaining the voluntary deep processing that amplifies germane learning. Together, these three frameworks specify the cognitive, instructional, and motivational mechanisms through which AI-assisted scaffolding is expected to improve both achievement and acceptability in combinatorics.

Meta-analytic evidence supports this approach. Ran et al. (2022) reported a weighted mean effect size of $g = 0.68$ for AI-based mathematics tutoring across 62 studies. Hillmayr et al. (2020) identified immediate feedback as the strongest moderator of digital tool effectiveness in secondary STEM across 105 studies. Crompton et al. (2022) found that step-by-step solution transparency is among the features most strongly associated with learning gains in AI applications. Etcuban and Pantinople (2018) demonstrated significant achievement gains when mobile applications were integrated into Philippine secondary mathematics. Hwang and Tu (2021) further confirmed that step-by-step feedback and adaptive sequencing are the most predictive features of AI-assisted mathematics learning outcomes.

Despite this growing evidence base, three gaps remain unaddressed in the literature. First, existing AI-in-mathematics research has concentrated on general-purpose intelligent tutoring systems or broad algebra and general-mathematics content (Hwang and Tu, 2021; Ran et al., 2022); to date, no published study has evaluated a mobile, photograph-based problem-solving application such as Gauthmath, despite its rapid uptake among Filipino learners. Second, combinatorics itself has received comparatively little attention in the AI-assisted learning literature relative to algebra or geometry, even though it is consistently identified as one of the most conceptually demanding MELC strands (Batanero et al., 1997; English, 2005; Gratela et al., 2023). Third, few studies situate an AI tool's evaluation within an explicit, integrated theoretical framework that jointly specifies the cognitive, developmental, and motivational mechanisms of change; most report outcomes without accounting for why an application should work in the first place. This study addresses these three gaps by providing, to the authors' knowledge, the first documented classroom evaluation of Gauthmath as a daily instructional supplement for combinatorics MELCs in a Philippine public secondary setting, using the integrated ZPD–CLT–SDT framework detailed above to render the mechanisms of scaffolding explicit and testable rather than assumed.

As a classroom teacher at Cotta National High School, the first author observed persistent learner difficulties with combinatorics during the third quarter of School Year (SY) 2024–2025. This action research investigated whether systematic, MELC-aligned Gauthmath integration

during the five-week combinatorics unit could produce measurable achievement gains and whether learners would find the application acceptable as a daily supplementary tool. The study represents a Technology–Mathematics strand integration within the Community and School STEM sub-theme: it applies an accessible, freely available AI technology to a persistent mathematics-achievement problem situated within a single public school community, reflecting STEM education's premise that disciplinary boundaries dissolve when a technological tool is deployed to address a genuine mathematics-learning need. Within this framing, the school functions as the immediate community in which the intervention was designed, delivered, and evaluated, and the findings are intended to inform both classroom practice and division-level decisions on technology-supported instruction, addressing the social and educational impact of persistent mathematics underperformance in Philippine secondary schools.

2. Method

2.1 Research Design

This study employed a one-group pre-test–post-test pre-experimental design (Campbell and Stanley, 1963), appropriate for classroom action research where random assignment is logistically infeasible. Represented as $O_1 X O_2$, where O_1 is the pre-test (Day 1), X is the five-week Gauthmath-supplemented combinatorics unit, and O_2 is the post-test (Day 25), the design tested whether a statistically significant performance difference existed before and after the intervention.

As a pre-experimental design, this approach carries recognised internal-validity limitations (Campbell and Stanley, 1963). Without a non-intervention comparison group, history (concurrent events or instruction outside the Gauthmath sessions), maturation (learners' ordinary cognitive or academic development across the five-week period), and testing effects (practice or sensitisation arising from completing structurally similar items at both administrations) cannot be fully disentangled from the effect of the intervention itself. Instrumentation was held constant, since the same researcher-made test was used at both administrations, and statistical regression toward the mean was anticipated given the purposive selection of low-scoring learners; this last threat was addressed analytically using the Lord–Novick (1968) correction described in Section 2.7. The remaining threats, history, maturation, and testing, were not statistically isolated in this design and are revisited as limitations in Section 3.

2.2 Research Locale and Participants

The study was conducted at Cotta National High School, Schools Division of Lucena City, Philippines, during the third quarter of SY 2024–2025. Participants were 30 Grade 10 learners enrolled in Mathematics, selected through purposive sampling. Inclusion criteria were: (a) Grade 10 enrolment; (b) pre-test score below the 40th percentile; (c) parental or guardian consent; and (d) learner assent. All 30 participants completed all 25 sessions (100% completion rate).

2.3 Intervention Procedure

The intervention spanned the five-week combinatorics MELC unit, with mathematics scheduled five days per week, yielding 25 sessions of 45 minutes each (1,125 minutes, equivalent to 18.75 hours total: $1,125 \div 60$). Sessions were conducted during regular class time and aligned directly with the prescribed DepEd MELC schedule: illustrating and solving permutations (Weeks 1–2), illustrating combinations and differentiating permutations from

combinations (Weeks 3–4), and solving integrated problems (Week 5). Table 1 presents the daily session schedule.

Each session followed a three-phase protocol: (a) a 15-minute independent problem attempt with Gauthmath closed; (b) a 20-minute application-assisted review using Gauthmath's step-by-step solutions; and (c) a 10-minute structured reflection via session log. Session logs were reviewed daily to identify persistent difficulties and adjust the following session's problem set. Live tutoring was disabled to ensure consistent automated scaffolding; learners accessed the application on personal smartphones ($n = 22$) or school-provided tablets ($n = 8$).

Table 1

Daily Session Schedule: Five-Week Combinatorics MELC Unit ($n = 30$, Grade 10)

Week	Day	MELC / Session Content	Gauthmath Problem Focus
1	1	Pre-test administration; orientation to Gauthmath	Baseline diagnostic
	2	Illustrates permutation: Fundamental Counting Principle and factorial notation	FCP basic cases; $n!$ notation
	3	Illustrates permutation: linear permutations (nPr): concept and formula	nPr formula; basic cases
	4	Illustrates permutation: linear permutations (nPr): applied problems	nPr applied problems
	5	Illustrates permutation: circular and distinguishable permutations	Circular P ; identical elements
2	6	Solves problems involving permutations: linear permutation problems	Mixed nPr problem sets
	7	Solves problems involving permutations: circular permutation problems	Mixed circular P problems
	8	Solves problems involving permutations: distinguishable permutation problems	Distinguishable P cases
	9	Solves problems involving permutations: mixed and applied problems	Mixed permutation sets
	10	Solves problems involving permutations: review and consolidation	Permutation review
3	11	Illustrates combination: concept of combination and nCr formula	nCr concept; basic cases
	12	Illustrates combination: nCr applied problems	nCr applied problems
	13	Illustrates combination: combinations with restrictions	Constrained nCr problems
	14	Differentiates permutation from combination of n objects taken r at a time: concept	P vs. C conceptual distinction
	15	Differentiates permutation from combination: applied problems	Mixed identification problems
4	16	Differentiates permutation from combination: classification problem sets	Mixed P and C classification
	17	Differentiates permutation from combination: complex justification problems	Reasoning and justification tasks
	18	Review and consolidation of Weeks 3–4 competencies	Combinations review
	19	Combinations: multi-step and complex problems	Multi-step nCr problems
	20	Formative assessment and targeted re-teaching based on session log analysis	Error-targeted problem sets

Week	Day	MELC / Session Content	Gauthmath Problem Focus
5	21	Solves problems involving permutations and combinations: integration, basic	Mixed P and C problems
	22	Solves problems involving permutations and combinations: applied real-world problems	Real-world P and C applications
	23	Solves problems involving permutations and combinations: complex multi-step problems	Multi-step mixed P and C
	24	Comprehensive review across all four combinatorics MELCs	Full-unit review
	25	Post-test administration	Summative assessment

Note. MELC = Most Essential Learning Competency (Department of Education [DepEd], 2020). Mathematics scheduled five days per week; each session was 45 minutes. Total: $25 \times 45 \text{ min} = 1,125 \text{ min} = 18.75 \text{ hours}$ ($1,125 \div 60$). Schedule aligned with the prescribed Q3 MELC calendar: permutations (Weeks 1–2), combinations and differentiation (Weeks 3–4), integration (Week 5).

2.4 Instruments

Achievement Test. A researcher-made 30-item multiple-choice achievement tests covered all four combinatorics MELCs, administered as both pre-test (Day 1) and post-test (Day 25). Item construction began with a two-way table of specifications (Table 2) that cross-referenced each MELC against Bloom's revised cognitive-process dimensions (Anderson and Krathwohl, 2001), yielding a target distribution of 18 procedural items (Remembering and Applying) and 12 conceptual items (Understanding, Analysing, and Evaluating). Each item was drafted in four-option multiple-choice format, with distractors constructed from the specific misconceptions documented in the combinatorics-reasoning literature, for example, confusing order-sensitive and order-insensitive counting situations, or omitting the fundamental counting principle in multi-step problems (Batanero et al., 1997; English, 2005). Draft items were reviewed by three subject-matter experts against three criteria, alignment with the target MELC, cognitive-level accuracy, and clarity of language, yielding an Item-level Content Validity Index (I-CVI) of .92 and a Scale-level CVI of .94, exceeding the .78 threshold (Polit and Beck, 2006); items flagged for ambiguity or misalignment were revised before piloting. The revised instrument was then piloted with 15 Grade 10 learners who did not participate in the main study, with item difficulty and discrimination indices computed to guide final refinement, followed by a second administration to the same pilot group to establish test-retest reliability ($r = .87$). $MPS = (\text{Mean Score} \div 30) \times 100$; mastery threshold: $MPS \geq 75\%$ (Department of Education, 2015).

Table 2

Test Blueprint: Item Distribution by MELC Competency and Cognitive Level

MELC Competency	Procedural	Conceptual	Total	%
Fundamental Counting Principle / Factorial	3	1	4	13.3
Permutations (Illustrates and Solves)	7	4	11	36.7
Combinations (Illustrates and Solves)	6	4	10	33.3
Differentiates P from C in Context	2	3	5	16.7
Total	18	12	30	100

Note. Competency labels aligned with DepEd (2020) prescribed MELC schedule. Taxonomy: Anderson and Krathwohl (2001).

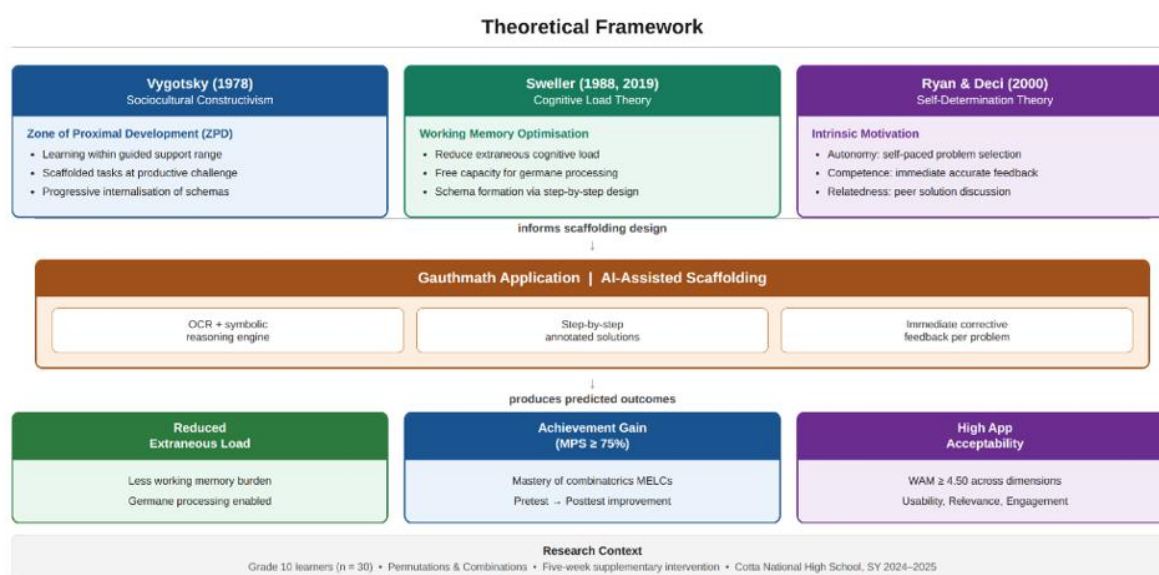
Acceptability Survey. A 15-item Likert-scale survey (1 = Strongly Disagree to 5 = Strongly Agree) assessed learner perceptions across usability (5 items), relevance (5 items), and engagement (5 items). Internal consistency: usability $\alpha = .89$; relevance $\alpha = .87$; engagement $\alpha = .91$, all exceeding the .70 threshold (Nunnally, 1978). Item-level CVI = .90. Responses were summarised using Weighted Average Mean (WAM): 4.50–5.00 = Strongly Agree; 3.50–4.49 = Agree; 2.50–3.49 = Neutral; 1.50–2.49 = Disagree; 1.00–1.49 = Strongly Disagree.

2.5 Conceptual Framework

Figures 1 and 2 present the theoretical and conceptual frameworks guiding this action research. The theoretical framework (Figure 1) illustrates how cognitive load theory, the Zone of Proximal Development, and self-determination theory converge on Gauthmath's scaffolding mechanisms to produce reduced extraneous load, achievement gain, and high acceptability. The conceptual framework (Figure 2) operationalises these into the study's input–process–output structure, connecting the five-week daily instruction process to the two measured output dimensions through three mediating variables.

Figure 1

Theoretical framework: cognitive load theory, Zone of Proximal Development, and self-determination theory as mechanisms for AI-assisted combinatorics scaffolding.

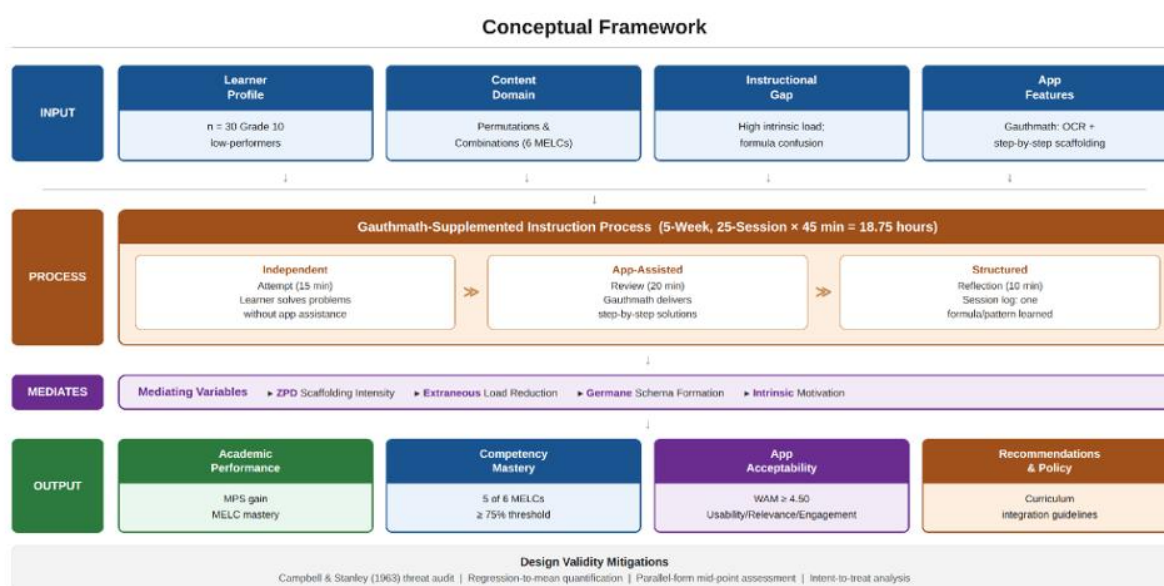


2.6 Data Collection

The pre-test was administered on Day 1 before any Gauthmath exposure to establish baseline competency profiles per MELC. Session logs were maintained daily throughout the five-week unit. The acceptability survey was administered immediately after Day 24. The post-test was administered on Day 25 under the same conditions as the pre-test.

Figure 2

Conceptual framework: input–process–output architecture with mediating variables.



2.7 Data Analysis

Analyses were performed in JASP. Descriptive statistics (mean, standard deviation, and MPS) were computed for pre-test and post-test. Item-level pass rates were computed per MELC. Acceptability data were summarised using WAM. A Shapiro–Wilk test on difference scores ($W = .961$, $p = .324$) confirmed normality, supporting a two-tailed dependent samples t -test at $\alpha = .05$. Effect size was calculated as Cohen’s d (Cohen, 1988). The Lord–Novick (1968) regression-to-mean formula was applied to estimate the artefactual share of the observed gain. Because sample size was fixed by the intact classroom section rather than determined through a priori power analysis, a post-hoc sensitivity check using G*Power (Faul et al., 2007) confirmed that the achieved sample ($n = 30$) provided power exceeding .999 to detect the observed effect at $\alpha = .05$, indicating the design was not underpowered despite the absence of a comparison group.

2.8 Ethical Considerations

Ethical clearance was obtained under DepEd Order No. 16, s. 2017 (Department of Education, 2017). Written informed consent was secured from all participants and parents or guardians; child assent forms were administered separately. Participants were informed of the voluntary nature of participation, the right to withdraw, data confidentiality, and the sole use of data for research. Data are stored in password-protected files scheduled for secure disposal five years after study completion.

Because the intervention required learners to submit photographs of their written work to a third-party commercial application, additional safeguards were applied specific to AI-assisted tool use. Gauthmath, like most commercial AI applications, collects account and usage information and processes data on servers outside the Philippines under its own privacy policy; this was disclosed to parents and guardians as part of the consent process, and learners were instructed to photograph only their mathematics work, avoiding the inclusion of names or other identifying information within the image itself wherever possible. Live tutoring, Gauthmath’s real-time human-assisted chat feature, was disabled for the duration of the study.

(Section 2.3), both to standardise the scaffolding learners received and to remove unsupervised interaction between learners and an external chat function. Use of the application was confined to the supervised 45-minute class period under teacher observation; learners were not required or encouraged to use Gauthmath outside class or on unmonitored personal accounts.

These safeguards align with the Data Privacy Act of 2012 (Republic Act No. 10173) (Republic of the Philippines, 2012), which governs the processing of personal information by third parties in the Philippines, and with DepEd Order No. 003, s. 2026, the Department's foundational guidelines on the use of artificial intelligence in basic education. That order was issued after this study's implementation but is consistent with the safeguards already described here, namely human oversight of AI-assisted instruction, transparency with learners and parents about when and how an AI tool is used, and protection of learner data (Department of Education, 2026). Consistent with these principles, the teacher-researcher retained full evaluative authority throughout the study; Gauthmath's outputs were treated strictly as a scaffolding aid for learners' own problem-solving rather than a substitute for it, and neither the achievement test nor the acceptability survey was processed by the application, since both instruments were administered independently of Gauthmath. Future implementations are encouraged to formalise these protections through school-managed accounts or offline alternatives where feasible, so that the benefits of AI-assisted scaffolding are not conditioned on learners' or families' independent data-privacy literacy.

3. Results and Discussion

Table 3 shows the item-level pre-test pass rates. The overall pre-test MPS was 27.44%, placing all 30 participants in the "Did Not Meet Expectations" category (Department of Education, 2015). No MELC competency reached the 75% mastery threshold. The most severely under-mastered competency was differentiating permutations from combinations in context (12.0%), consistent with Batanero et al.'s (1997) implicit model confusion thesis and English's (2005) finding that combinatorics difficulty stems from abstract enumeration demands exceeding working memory capacity. This diagnostic profile directly shaped the session sequence in Table 1.

Table 3

Item Pass Rates by MELC Competency: Pre-test and Post-test

MELC Competency	Pre %	Post %	Gain (pp)	Mastery
Fundamental Counting Principle	41.7	91.7	+50.0	Yes
Permutations (Illustrates and Solves)	33.3	85.0	+51.7	Yes
Circular and Distinguishable Permutation	17.3	73.3	+56.0	Approaching
Combinations (Illustrates and Solves)	30.0	88.3	+58.3	Yes
Combinations with Restrictions	19.2	79.2	+60.0	Yes
Differentiates P from C in Context	12.0	76.0	+64.0	Yes
Overall	27.44	81.56	+54.12	Mastered

Note. Mastery: MPS $\geq$ 75% (Department of Education, 2015). pp = percentage points.

Table 4 presents learner acceptability ratings. Usability (WAM = 4.58, α = .89), relevance (WAM = 4.63, α = .87), and engagement (WAM = 4.69, α = .91) were all interpreted as Strongly Agree. The highest-rated single item was supplementary value relative to classroom activities

(WAM = 4.93), confirming that learners perceived Gauthmath as extending rather than duplicating their daily lessons. This is precisely the ZPD-aligned supplementary role intended. The progressive usability → relevance → engagement gradient is theoretically coherent with Ryan and Deci's (2000) account of intrinsic motivation: the application crossed the threshold from utility to motivation, producing the voluntary, deep processing that amplifies germane schema formation (Sweller et al., 2019).

Table 4

Summary of Learner Acceptability Across Dimensions (n = 30)

Dimension	WAM	α	Interpretation
Usability	4.58	.89	Strongly Agree
Relevance	4.63	.87	Strongly Agree
Engagement	4.69	.91	Strongly Agree
Overall	4.63	.89	Strongly Agree

Note. Overall α computed across all 15 items. Scale: 4.50–5.00 = Strongly Agree.

Table 5 presents the performance statistics. The dependent samples *t*-test yielded $t(29) = -22.4$, $p < .001$, $d = 4.09$, 95% CI $[-17.7, -14.8]$, confirming a statistically significant and practically large gain. The Lord–Novick (1968) regression-to-mean estimate attributes approximately 24% of the raw gain to statistical artefact; the conservative residual gain of approximately 12.31 points (approximately 41 percentage points in MPS terms) still substantially exceeds the AI tutoring meta-analytic mean of $g = 0.68$ (Ran et al., 2022). These results are coherent across all three theoretical levels: daily step-by-step scaffolding reduced extraneous cognitive load (CLT), ZPD-aligned feedback enabled progressive schema internalisation across 25 sessions, and high motivational engagement (SDT, WAM = 4.69) amplified germane processing.

Table 5

Descriptive Statistics and t-Test Results: Pre-test and Post-test

	<i>n</i>	<i>M</i>	<i>SD</i>	MPS (%)	Level	<i>t</i> (<i>df</i> = 29)	<i>d</i>
Pre-test	30	8.23	2.28	27.44	DNM		
Post-test	30	24.47	2.73	81.56	Mastered	-22.4***	4.09

Note. *** $p < .001$ (two-tailed). DNM = Did Not Meet Expectations (MPS < 75%). 95% CI $[-17.7, -14.8]$. Cohen (1988): large $d \geq .80$.

3.1 Interpretation of Acceptability Findings

The consistently high acceptability ratings, particularly the near-ceiling rating for supplementary value (WAM = 4.93), plausibly reflect more than surface-level enjoyment. Because Gauthmath returned a fully worked, step-by-step solution immediately after each independent attempt, learners received the kind of prompt, error-specific feedback that is rarely available to an individual learner in a class of thirty during a single 45-minute period; this addresses a structural constraint of the traditional classroom rather than merely adding novelty. The three-phase protocol also preserved learner agency, since the independent attempt always preceded application use, positioning Gauthmath as a check on the learner's

own reasoning rather than a replacement for it. This sequencing is consistent with the self-determination-theoretic account (Ryan and Deci, 2000) advanced in Section 1: learners were not passively receiving answers but actively testing their own solutions against expert-level feedback, which plausibly sustained the sense of competence and autonomy reflected in the acceptability ratings.

3.2 Practical Implementation Challenges

Despite these gains, several practical implementation challenges warrant discussion for teachers considering similar integration. Device access was not uniform: 22 learners used personal smartphones and 8 relied on school-provided tablets, and classroom connectivity occasionally slowed the application-assisted review phase, requiring the teacher-researcher to stagger submissions or briefly extend the 20-minute review window; schools without a comparable minimum of shared devices or reliable connectivity would need to budget additional time or resources before adopting this model. Reliance on a single commercial application also introduces a sustainability risk, since Gauthmath's feature set, pricing, or availability could change independently of the school's instructional planning; teachers adopting this approach should treat the specific application as replaceable within the underlying three-phase pedagogical structure rather than as an indispensable component. Sustained daily use further raises a legitimate question of learner dependency, that is, whether learners increasingly defer to the application's worked solutions rather than attempting problems independently. The fixed sequencing of the three-phase protocol, with independent attempt strictly preceding application access, was designed specifically to guard against this, and the sustained gains through Week 5 (Table 3) suggest that independent competency did develop rather than being masked by application use; nonetheless, this study did not include a delayed post-test without application access, and future implementations should monitor for over-reliance directly, for instance by periodically assessing performance under application-free conditions. Finally, integrating a new digital routine into daily instruction required an initial adjustment period for classroom management, including distributing and collecting devices and managing screen time within the 45-minute period, which teachers replicating this model should plan for in the first one to two sessions rather than treating the protocol as immediately turnkey.

3.3 Methodological Limitations

These findings should also be interpreted in light of the design limitations foreshadowed in Section 2.1. As a one-group pre-test–post-test design without a comparison group, the study cannot fully rule out history (concurrent exposure to combinatorics content through means other than the Gauthmath sessions, such as homework or peer study), maturation (learners' general cognitive or academic development across the five-week period), or testing effects (familiarity with the item format gained simply by completing the pre-test) as partial contributors to the observed gain, alongside statistical regression to the mean, which was addressed directly through the Lord–Novick (1968) correction above. The magnitude of the residual, artefact-adjusted gain, approximately 41 percentage points, and its coherence with the theoretical mechanisms specified in Section 1 make a purely artefactual explanation unlikely, but a quasi-experimental design with a matched comparison group, as recommended in Section 4, remains necessary to isolate the intervention's causal contribution with confidence. The single-school, single-section sample ($n = 30$) further limits generalisability to other MELC strands, grade levels, or school contexts with different resource profiles.

4. Conclusion

This action research demonstrated that integrating Gauthmath as a daily AI-assisted supplement throughout the five-week combinatorics MELC unit produced both strong learner acceptability (Overall WAM = 4.63, Strongly Agree) and a statistically significant, practically large achievement gain (MPS from 27.44% to 81.56%, $d=4.09$). Five of six MELC competency clusters reached mastery; circular and distinguishable permutations (73.3%) remain the only area below the 75% threshold, identifying a clear target for continued instructional attention. From a practitioner perspective, the three-phase session protocol integrated into regular 45-minute class periods across the full MELC unit, using devices already available to learners and the school. This offers a practical, low-cost model for other mathematics teachers to adapt, provided the device, connectivity, and classroom-management considerations discussed in Section 3.2 are planned for in advance. The Technology–Mathematics STEM integration demonstrated here provides a concrete, MELC-aligned model for the Community and School sub-theme.

The study's primary contribution is providing, to the authors' knowledge, the first documented classroom evaluation of Gauthmath targeting combinatorics MELCs in the Philippine secondary context, grounded in an integrated triple-theoretical framework that makes the mechanisms of improvement explicit rather than assumed. Future research should employ a quasi-experimental design with a matched comparison group to strengthen causal attribution, include a delayed post-test administered without application access to assess retention and guard against over-reliance, and extend the model to other high-intrinsic-load MELC strands such as probability and statistics. Division offices should ensure equitable device and connectivity access so that AI-assisted learning benefits are not confined to better-resourced schools.

References

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
https://openlibrary.org/books/OL18156636M/A_taxonomy_for_learning_teaching_and_assessing
- Ando, K. N., Basilisco, J. S., Deniega, A. J., Gador, K. L., Geraldo, P. J. A., Gipulao, W. E. M. M., Redondo, M. R., Sumampong, M. M. S., Talledo, M. T., Villacis, F. T., Taneo, J. D., Cabello, C. A., & Minyamin, A. V. (2022). Learning without learning in the new normal: College education students' lived experiences in blended learning modality. *Psychology and Education: A Multidisciplinary Journal*, 2(6), 455–464.
<https://ejournals.ph/article.php?id=20388>
- Batanero, C., Navarro-Pelayo, V., & Godino, J. D. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics*, 32(2), 181–199. <https://doi.org/10.1023/A:1002954428327>
- Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, 87(2), 309–344.
<https://doi.org/10.3102/0034654316670999>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally. <https://search.worldcat.org/title/7640465>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
<https://www.routledge.com/Statistical-Power-Analysis-for-the-Behavioral-Sciences/Cohen/p/book/9780805802832>

- Crompton, H., Jones, M. V., & Burke, D. (2022). Affordances and challenges of artificial intelligence in K-12 education: A systematic review. *Journal of Research on Technology in Education*, 56(3), 248–268. <https://doi.org/10.1080/15391523.2022.2121344>
- Department of Education. (2015). *Policy guidelines on classroom assessment for the K to 12 basic education program* (DepEd Order No. 8, s. 2015). <https://www.deped.gov.ph/2015/04/01/do-8-s-2015-policy-guidelines-on-classroom-assessment-for-the-k-to-12-basic-education-program/>
- Department of Education. (2017). *Research management guidelines* (DepEd Order No. 16, s. 2017). <https://www.deped.gov.ph/2017/03/20/do-16-s-2017-research-management-guidelines/>
- Department of Education. (2019). *PISA 2018 Philippine national report*. <https://www.deped.gov.ph/publications/pisa-2018-philippine-national-report/>
- Department of Education. (2020). *Guidelines on the use of the Most Essential Learning Competencies (MELCs)* (Annex C of DepEd Order No. 12, s. 2020). <https://www.deped.gov.ph/2020/06/19/june-19-2020-do-012-2020-adoption-of-the-basic-education-learning-continuity-plan-for-school-year-2020-2021-in-the-light-of-the-covid-19-public-health-emergency/>
- Department of Education. (2026). *Foundational guidelines on artificial intelligence (AI) in basic education* (DepEd Order No. 003, s. 2026). <https://www.deped.gov.ph/2026/02/20/february-20-2026-do-003-s-2026-foundational-guidelines-on-artificial-intelligence-ai-in-basic-education/>
- English, L. D. (2005). Combinatorics and the development of children's combinatorial reasoning. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 121–141). Springer. https://doi.org/10.1007/0-387-24530-8_6
- Etcuban, J. O., & Pantinople, L. D. (2018). The effects of mobile application in teaching high school mathematics. *International Electronic Journal of Mathematics Education*, 13(3), 249–259. <https://doi.org/10.12973/iejme/3906>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Gratela, K. F., Dio, R. V., & Deri, R. A. (2023). Mapping of Philippine Kto10 most essential learning competencies (MELCs) in probability and statistics. *EDUCATUM Journal of Science, Mathematics and Technology*, 10(2), 19–28. <https://doi.org/10.37134/ejsmt.vol10.2.3.2023>
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153, 103897. <https://doi.org/10.1016/j.compedu.2020.103897>
- Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, 9(6), 584. <https://doi.org/10.3390/math9060584>
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley Publishing Company. <https://search.worldcat.org/title/222679>
- Mishra, L., Gupta, T., & Shree, A. (2020). Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *International Journal of Educational Research Open*, 1, Article 100012. <https://doi.org/10.1016/j.ijedro.2020.100012>

- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
<https://search.worldcat.org/title/3167590>
- OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Ran, H., Kim, N. J., & Secada, W. G. (2022). A meta-analysis on the effects of technology's functions and roles on students' mathematics achievement in K-12 classrooms. *Journal of Computer Assisted Learning*, 38(1), 258–284.
<https://doi.org/10.1111/jcal.12611>
- Republic of the Philippines. (2012). *Republic Act No. 10173: Data Privacy Act of 2012*.
<https://privacy.gov.ph/data-privacy-act/>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261–292.
<https://doi.org/10.1007/s10648-019-09465-5>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

STORIES TO EXPERIMENT: THE STREAM STORY CYCLE FOR COMMUNICATIVE ESL LEARNING

Hjh Nor Nasriah Akmal binti Haji Kasim

Sekolah Rendah OKSB Kilanas, Ministry of Education, Brunei Darussalam

<Nasriah.kasim@hotmail.com>

ABSTRACT

In many English as a Second Language (ESL) classrooms across Southeast Asia, language instruction is often separated from meaningful problem-solving activities. Traditional lessons emphasise grammar drills and isolated vocabulary practice, while STREAM (Science, Technology, Reading, Engineering, Arts, and Mathematics) activities may prioritise construction without explicitly supporting language development. This paper introduces the STREAM Story Cycle, a pedagogical framework that integrates storytelling, collaborative design inquiry, and structured language reflection to support communicative language development. Drawing on narrative cognition, constructionism, social constructivism, and interactionist theory, the framework positions stories as catalysts for collaborative design challenges. A mixed-methods classroom study with Year 5 ESL students in Brunei Darussalam shows that while STREAM-based learning supports vocabulary exposure and engagement across all learners, differences emerge in confidence and productive language use, particularly among lower-achieving students. The findings highlight the importance of combining meaningful learning contexts with explicit language scaffolding. Beyond Brunei, the STREAM Story Cycle offers a culturally adaptable model for Southeast Asian classrooms, where storytelling traditions are rich and diverse, aligning with SEAMEO's regional priorities of fostering communicative competence and 21st-century skills.

Keywords: STREAM Education, ESL Pedagogy, Storytelling in Language Learning, Communicative Competence, Collaborative Inquiry, Primary Education

1. Introduction: The Need for Narrative Inquiry

In many English as a Second Language (ESL) contexts, language instruction remains siloed, focusing on isolated grammar points and vocabulary drills that offer limited opportunities for meaningful communication. Students may successfully complete written exercises yet struggle to use language authentically in communicative situations.

On the other hand, when Science, Technology, Reading, Engineering, Arts and Mathematics (STREAM) activities are introduced in classrooms, the emphasis often shifts heavily toward experimentation and construction. While these activities promote engagement and creativity, language learning may become secondary to the building process. As a result, opportunities for structured language development are sometimes overlooked.

For effective language acquisition to occur, learners require opportunities to use language in meaningful and purposeful contexts. Learners must not only encounter vocabulary but also use language to explain ideas, negotiate solutions, and reflect on their thinking. Without authentic contexts for communication, student engagement may decline and language retention can remain limited.

Stories provide an emotional and narrative framework that captures learners' attention (Bruner, 1991), while inquiry-based design tasks offer the cognitive challenge needed to sustain engagement. When stories introduce problems that require collaborative solutions, language becomes a natural tool for communication rather than a purely academic exercise.

This paper proposes the STREAM Story Cycle, a teaching framework that integrates storytelling, collaborative design inquiry, and language reflection to support communicative language development in ESL classrooms.

2. Literature Review: The Pillars of the Framework

The STREAM Story Cycle draws upon several complementary theoretical perspectives from education and second language acquisition research. Together, these perspectives highlight the importance of narrative context, collaborative inquiry, and meaningful interaction in supporting language development.

2.1 Storytelling: Narrative Cognition and Meaning Making

Humans naturally make sense of the world through stories. Narrative structures help learners organise complex information into meaningful sequences of events.

Bruner (1991) argues that narrative thinking allows individuals to construct meaning by arranging experiences into coherent story structures. For language learners, stories provide a familiar cognitive framework that reduces cognitive load, enabling students to focus on understanding and using new vocabulary within meaningful contexts.

Within the STREAM Story Cycle, storytelling functions as the entry point of the learning experience. Narrative contexts introduce vocabulary, establish problems, and motivate learners to engage in inquiry. By embedding learning challenges within stories, students are more likely to remember both the events of the narrative and the language associated with them.

2.2 STREAM Learning: Constructionism and the Affective Filter

Hands-on learning plays a significant role in promoting engagement and reducing language anxiety. Seymour Papert's theory of constructionism suggests that learning occurs most effectively when learners actively construct tangible artefacts that represent their thinking (Papert & Harel, 1991).

In the STREAM Story Cycle, learners build physical solutions to narrative problems. These construction tasks shift students' attention away from speaking English perfectly and toward solving the design challenge. This change in focus helps lower the affective filter described by Stephen Krashen, which refers to emotional barriers that can prevent language input from being processed effectively (Krashen, 1982).

When students collaborate on building tasks, language emerges naturally as they explain ideas, request materials, and discuss improvements. The design process therefore becomes a meaningful context for communicative language use.

2.3 Collaborative Learning: Social Constructivism

Collaborative learning environments encourage learners to develop language through interaction with peers. Vygotsky (1978) proposed that learning occurs most effectively within the Zone of Proximal Development (ZPD), where learners can perform tasks with support from more capable peers.

During the collaborative planning and construction stages of the STREAM Story Cycle, students negotiate ideas, explain reasoning, and clarify misunderstandings. These interactions encourage learners to reformulate language in ways that support shared understanding. Through these interactions, language becomes a tool for collaborative thinking rather than simply an object of study.

2.4 Interactionist Theory: Negotiation of Meaning

The interactionist perspective on second language acquisition highlights the role of communication breakdowns in language learning. Long's (1996) Interaction Hypothesis suggests that language development occurs when learners negotiate meaning during communication.

Collaborative design activities naturally generate moments where learners must clarify instructions, adjust explanations, or ask for clarification. These interactions encourage deeper processing of language input and promote language acquisition.

The STREAM Story Cycle intentionally creates opportunities for these communicative interactions by embedding language use within collaborative problem-solving tasks.

2.5 Contemporary Perspectives on STREAM, Storytelling and Communicative Language Learning

Recent research demonstrates growing interest in integrating interdisciplinary learning with language education to promote both academic achievement and twenty-first-century competencies. Studies on STREAM education consistently report improvements in learners' creativity, collaboration, critical thinking and problem-solving through authentic inquiry-based experiences. However, language development is frequently reported as an incidental outcome rather than an explicit instructional objective, highlighting the need for frameworks that intentionally integrate communicative language learning within STREAM lessons (Perignat & Katz-Buonincontro, 2019; Gatcho, 2024).

Similarly, contemporary storytelling research has evolved beyond traditional read-aloud activities towards collaborative and digital storytelling approaches. Recent systematic reviews and meta-analyses indicate that storytelling significantly enhances vocabulary acquisition, learner motivation, speaking confidence and overall language proficiency by providing meaningful contexts for authentic communication (Lee, 2022; Lim et al., 2022, Pérez Agustín & de-la-Peña Álvarez, 2025).

Communicative language teaching likewise emphasises authentic interaction, collaboration and purposeful language use rather than isolated grammar instruction (Richards, 2006). Although inquiry-based learning naturally creates opportunities for negotiation of meaning, relatively few instructional models explicitly scaffold vocabulary learning, oral explanation and reflective language production throughout collaborative engineering design. Consequently, there remains a need for pedagogical frameworks that intentionally integrate storytelling, inquiry and communicative language development within a coherent instructional cycle.

3. Research Gap and Scholarly Contribution

Although storytelling, communicative language teaching and STREAM education have each been widely investigated, these approaches have largely developed as separate strands of educational research. Storytelling research primarily focuses on reading comprehension, vocabulary acquisition and oral language development (Lee, 2022), whereas STREAM education emphasises interdisciplinary inquiry, creativity and engineering design (Perignat & Katz-Buonincontro, 2019). Meanwhile, communicative language teaching promotes meaningful interaction but provides limited guidance on embedding language instruction within engineering inquiry.

Consequently, few instructional frameworks explicitly integrate narrative engagement, collaborative engineering inquiry and structured language reflection within a single pedagogical model. Existing approaches often assume that communication will emerge naturally during collaborative activities but provide limited guidance regarding how vocabulary development, oral explanation and reflective language production should be intentionally scaffolded throughout the inquiry process.

The STREAM Story Cycle addresses this gap by proposing a recursive instructional framework that embeds language learning throughout every stage of engineering inquiry. Rather than positioning communication as an incidental outcome of collaboration, the framework deliberately integrates contextualised vocabulary learning, collaborative planning, structured oral explanation and narrative reconstruction as core instructional components. In doing so, it positions communicative competence as a central learning outcome alongside engineering design and interdisciplinary problem-solving.

Scholarly Contribution of the STREAM Story Cycle

The principal contribution of the STREAM Story Cycle extends beyond combining storytelling with STREAM education. Previous studies have demonstrated the effectiveness of storytelling in supporting language learning and the value of STREAM in fostering inquiry, creativity and collaboration. However, these approaches have generally been implemented independently or with limited attention to explicit communicative language development (Lee, 2022; Gatcho, 2024).

The STREAM Story Cycle contributes to the literature in three important ways.

First, it reconceptualises storytelling as the catalyst for engineering inquiry rather than merely as a literacy activity. Stories introduce authentic problems that motivate learners to investigate, design and communicate possible solutions.

Second, the framework intentionally embeds language scaffolding throughout the engineering design process. Unlike many inquiry-based approaches where communication is expected to emerge naturally, each stage explicitly promotes vocabulary acquisition, negotiation of meaning, oral explanation and academic language production.

Third, the framework culminates in Story Reconstruction, requiring learners to synthesise engineering knowledge and language learning by retelling or rewriting narratives using newly acquired vocabulary and design experiences. This recursive process strengthens conceptual understanding while reinforcing communicative competence.

Collectively, these features distinguish the STREAM Story Cycle from existing storytelling, inquiry-based and STREAM approaches by positioning language development as an explicit instructional objective rather than a secondary outcome of interdisciplinary learning.

4. Methodology

4.1 Research Design

This study employed a mixed-methods design, combining quantitative survey data with qualitative interview data. This approach provided measurable evidence of vocabulary learning, confidence, and engagement, while also capturing student voice and deeper insights into their experiences.

The integration of both methods enabled data triangulation, enhancing the validity and reliability of the findings.

4.2 Participants

The participants were 25 Year 5 ESL students from a government primary school in Brunei Darussalam who were engaged in STREAM-based lessons. The class consisted of:

- 6 High Achievers (HA)
- 12 Middle Achievers (MA)
- 7 Low Achievers (LA)

All students completed the survey. A purposive sample of six students (2 HA, 2 MA, 2 LA) participated in semi-structured interviews to ensure representation across achievement levels.

4.3 Instruments

4.3.1 Survey (Quantitative)

A Likert-scale survey (1 = Strongly Disagree to 5 = Strongly Agree) was administered, consisting of four sections:

- Vocabulary learning
- STREAM learning experience
- Confidence in English
- Engagement

Open-ended items were included to capture brief reflections. The survey was aligned with lesson objectives and reviewed for clarity and age-appropriateness.

4.3.2 Interviews (Qualitative)

Semi-structured interviews were conducted using child-friendly prompts focusing on:

- Story connection
- Vocabulary learning
- STREAM activity
- Language use
- Reflection

The format allowed flexibility while maintaining consistency across participants.

4.3.3 Procedure

All students completed the survey immediately after the STREAM lesson. Responses were collected anonymously to encourage honest feedback.

Six students participated in interviews lasting 10–15 minutes. Interviews were audio-recorded, transcribed, and analysed thematically.

4.3.4 Data Analysis

Quantitative data were analysed using descriptive statistics (percentages and frequency distributions), with additional comparison across HA, MA, and LA groups.

Qualitative data were analysed using inductive thematic analysis, identifying recurring themes and selecting representative quotes.

4.3.5 Ethical Considerations

Parental consent and student assent were obtained prior to data collection. Participation was voluntary, and anonymity and confidentiality were maintained.

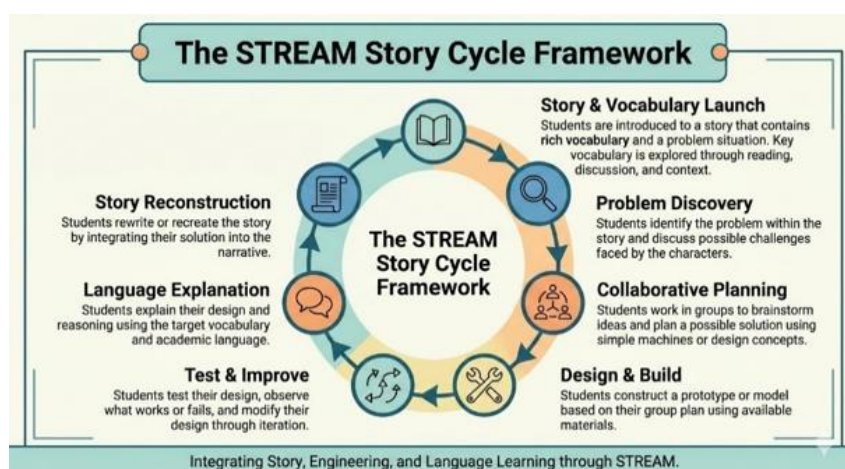
5. The STREAM Story Cycle Framework

The STREAM Story Cycle is designed as a recursive learning model in which narrative engagement, collaborative inquiry, and language reflection interact to support both conceptual understanding and communicative language development.

The framework follows a progression from narrative engagement to collaborative design inquiry and finally to structured language reflection. Each stage contributes to language development in a distinct way: narrative contexts introduce vocabulary and motivation, collaborative design tasks generate opportunities for interaction, and reflective language activities consolidate learning. The stages of the STREAM Story Cycle are illustrated in Figure 1.

Figure 1

The STREAM Story Cycle Framework.



Note. Developed by the author based on Bruner (1991), Vygotsky (1978), Papert and Harel (1991), and Long (1996).

5.1 The Seven Stages of the STREAM Story Cycle

Stage 1: Story and Vocabulary Launch

A narrative is introduced that presents a problem and key vocabulary.

Stage 2: Problem Discovery

Students identify and discuss the challenge presented in the story.

Stage 3: Collaborative Planning

Small groups brainstorm possible solutions and plan their designs.

Stage 4: Design and Build

Students construct physical prototypes that address the problem.

Stage 5: Test and Improve

Designs are tested and refined through evaluation and modification.

Stage 6: Language Explanation

Students articulate their reasoning while teachers support the development of precise language structures.

Stage 7: Story Reconstruction

Students retell or rewrite the story incorporating their engineered solutions.

Table 1

Phases of the Learning Cycle

Phase	Stages	Learning Purpose
Narrative	Story Launch, Problem	Contextualised vocabulary and motivation
Engagement	Discovery	
Inquiry and Construction	Planning, Build, Test	Collaboration and negotiation of meaning
Language Reflection	Explanation, Reconstruction	Language consolidation and narrative production

Note. Developed by the author based on Bruner (1991), Vygotsky (1978), Papert and Harel (1991), and Long (1996).

5.2 Integration of Literacy and Aesthetics (Reading & Arts)

While the framework heavily leverages engineering inquiry and communicative output, it relies equally on reading literacy and artistic expression to complete the learning cycle. The "Reading" component extends beyond simple passive listening during the Story Launch; it involves active textual decoding, critical comprehension strategies, and the interrogation of narrative plot points to uncover embedded design flaws. Concurrently, the "Arts" are explicitly integrated during the Collaborative Planning and Design phases. Students engage in visual arts through schematic drafting and prototype ideation, and performative arts during the final presentation stages, using spatial, structural, and aesthetic design choices to communicate their technical reasoning.

6. Classroom Implementation Examples

6.1 Implementation Phases

Example 1: Three Billy Goats Gruff Bridge Challenge

A lesson based on the traditional story *The Three Billy Goats Gruff* demonstrates how the STREAM Story Cycle can be implemented in an ESL classroom.

Stage 1: Story Context

Three goats must cross a bridge guarded by a troll.

Stage 2: Problem Discovery

Students discover that the bridge is unstable and cannot safely support the goats.

Stage 3: Collaborative Planning

Students discuss possible materials and structures that could strengthen the bridge.

Stage 4: Design and Build

Groups construct bridges using materials such as paper, sticks, and tape.

Stage 5: Test and Improve

The bridges are tested using weighted objects representing the troll.

Stage 6: Language Explanation

Students explain their design choices using casual language such as "*because*" and "*so that*".

Example student explanations include:

"We made the bridge wider because the troll is heavy."

"We added more sticks so that the bridge will not break."

Stage 7: Story Reconstruction

Students rewrite the ending of the story describing how their improved bridge allows the goats to cross safely.

Example 2: Tasek Merimbun Engineering Challenge

The STREAM Story Cycle can also be adapted to local cultural contexts. The following example draws upon the Bruneian folk narrative associated with Tasek Merimbun.

1. Narrative Hook

In the old days, the people of Merimbun wanted to reach the moon to capture its silver light for the lake. However, their ladders always wobbled and collapsed in the strong winds.

Students are invited to become members of the Royal Engineering Guild, tasked with designing a modular ladder capable of withstanding strong winds while supporting the weight of the moon.

2. Design Challenge

Students must construct a Modular Extension Ladder capable of withstanding a simulated wind test using a fan while supporting a weighted object representing the moon.

Table 2*Materials Needed to Create a Ladder*

Material	Quantity	Engineering Purpose
Ice cream sticks	30	Structural beams
Masking tape	1 metre	Joint reinforcement
Cardboard	2 pieces	Base foundation
Pipe cleaners	5	Cross bracing

Note. Materials were selected based on availability and suitability for the classroom engineering task.

6.2 Testing Phase

6.2.1 Students Test Their Ladders Using Two Criteria:

Wind Test – placing the ladder in front of a fan to simulate strong wind.

Load Test – placing a weighted object representing the moon at the top of the ladder.

6.2.2 Iteration Log

Students record one element of their design they kept and one change they made after testing.

6.2.3 Language Scaffolding

Sentence frames support students in articulating their reasoning:

"We chose a wide base so that the ladder would not fall."

"We added triangular supports because triangles are strong."

"Our second design was stronger than the first because we reinforced the joints."

6.3 Differentiated Story Reconstruction Outcomes

The final narrative rewrite serves as a formative assessment tool for assessing student language consolidation. High-achieving students frequently produced expanded, independent paragraphs, integrating complex conditional clauses (e.g., *"If the Royal Guild had not reinforced the joints with cross bracing, the ladder would have succumbed to the Merimbun gale"*). Conversely, lower-achieving students successfully reconstructed the narrative ending by utilising target sentence frames as structural anchors. A representative sample from a low-achieving group demonstrated functional communicative competence by writing: *"The engineers built a new ladder. We chose a wide base so that the ladder would not fall, and the people finally caught the moon's light"*. This demonstrates that the reconstruction phase accommodates varying degrees of linguistic autonomy while ensuring all tier levels manipulate the target vocabulary meaningfully.

7. Findings and Observations

7.1 Vocabulary Learning

Table 3

Vocabulary Learning Outcomes by Achievement Level

Achievement Group	Vocabulary Learned (%)	Vocabulary Used Correctly (%)
High Achievers (HA)	90	85
Middle Achievers (MA)	70	60
Low Achievers (LA)	50	35

Note. Vocabulary recognition and productive vocabulary use following implementation of the STREAM Story Cycle.

7.2 STREAM Learning Experience

Table 4

Students' STREAM Learning Experience by Achievement Level

Achievement Group	Story Supported Understanding (%)	Activity Supported Understanding (%)
High Achievers (HA)	95	90
Middle Achievers (MA)	75	70
Low Achievers (LA)	55	45

Note. Students' perceptions of how storytelling and STREAM activities supported learning.

The story functioned as a scaffold, helping students conceptualise the task, although its impact was reduced among lower-achieving learners.

7.3 Confidence in English

Table 5

Students' Confidence in Using English by Achievement Level

Achievement Group	Confidence (%)	Used Vocabulary During Activities (%)
High Achievers (HA)	80	85
Middle Achievers (MA)	55	60
Low Achievers (LA)	30	40

Note. Self-reported confidence in communicating in English after participation.

7.4 Engagement

Table 6

Student Engagement During the STREAM Story Cycle

Achievement Group	Enjoyment (%)
High Achievers (HA)	95
Middle Achievers (MA)	90
Low Achievers (LA)	80

Note. Representative student comment: "This is fun because we can do and not just write."

7.5 Integrated Findings

Overall, STREAM-based learning supported engagement and vocabulary exposure across all learners. However, disaggregated data revealed that lower-achieving students demonstrated engagement without equivalent gains in confidence or productive language use.

8. Discussion

As shown in Table 3, students across all achievement levels demonstrated positive vocabulary learning outcomes following implementation of the STREAM Story Cycle. High-achieving students recorded the highest levels of vocabulary recognition (90%) and productive vocabulary use (85%), whereas low-achieving students demonstrated lower vocabulary recognition (50%) and productive use (35%). These findings indicate that the framework supported vocabulary acquisition across all learner groups, although productive language use remained more challenging for lower-achieving students. This pattern is consistent with Nation's (2001) distinction between receptive and productive vocabulary knowledge.

The findings presented in Table 4 indicate that storytelling played an important role in supporting students' understanding of the engineering tasks. Students reported that both the narrative context and the STREAM activities enhanced comprehension, although the benefits were greater among high-achieving learners. These findings support Bruner's (1991) argument that narratives provide cognitive structures that help learners organise and interpret new experiences.

Results presented in Table 5 reveal substantial differences in students' confidence when communicating in English. Although high-achieving students reported relatively high confidence (80%), confidence among low-achieving students remained considerably lower (30%). This suggests that meaningful contexts alone are insufficient to promote communicative competence and that explicit language scaffolding remains necessary. The findings support Swain's (1985) Output Hypothesis, which argues that learners require opportunities for structured language production.

The consistently high engagement levels reported in Table 6 suggest that the STREAM Story Cycle successfully motivated students regardless of achievement level. Even low-achieving learners reported high enjoyment (80%), indicating that the integration of storytelling and hands-on design activities created an inclusive and motivating learning environment. These findings align with Papert and Harel's (1991) constructionist perspective and Krashen's (1982)

Affective Filter Hypothesis, which propose that meaningful, low-anxiety learning environments facilitate language acquisition.

Finally, classroom observations and interview responses indicate that collaborative planning and design activities encouraged students to explain ideas, negotiate solutions, and clarify misunderstandings during group work. These communicative interactions are consistent with Vygotsky's (1978) theory of social constructivism and Long's (1996) Interaction Hypothesis, both of which emphasise the importance of social interaction and negotiation of meaning in second language development.

The findings suggest that communicative language development is strengthened when language instruction is intentionally embedded within collaborative inquiry rather than taught separately from authentic learning experiences. The STREAM Story Cycle therefore extends existing theories of narrative learning (Bruner, 1991), social constructivism (Vygotsky, 1978), constructionism (Papert & Harel, 1991) and interactionist theory (Long, 1996) by demonstrating how storytelling, engineering design and reflective language production can function as mutually reinforcing components within a single instructional framework. Rather than positioning language as a support for inquiry, the framework conceptualises communication as the mechanism through which inquiry is negotiated, explained and reconstructed.

9. Conclusion

The STREAM Story Cycle offers a practical and culturally responsive framework for integrating language learning with interdisciplinary inquiry. By combining storytelling, collaborative design, and reflective language use, it creates meaningful opportunities for communicative competence.

Findings show that while STREAM-based learning promotes engagement and contextualised vocabulary acquisition, additional scaffolding is needed for productive language use. Importantly, the framework is adaptable to Southeast Asian contexts, where local narratives can serve as powerful entry points for inquiry-based language learning.

This positions the STREAM Story Cycle as a scalable approach aligned with SEAMEO's priorities of multilingual education, inclusive pedagogy, and 21st-century skill development. Future research should explore long-term retention, cross-country applications, and integration into teacher professional development programmes to strengthen ESL pedagogy across ASEAN classrooms.

References

- Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, 18(1), 1–21.
<https://doi.org/10.1086/448619>
- Gatcho, A. R. G. (2024). Streaming media for English language learners' vocabulary development: A systematic review of features and challenges. *International Journal of Learning, Teaching and Educational Research*, 23(3), 448–468.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.
https://sdrashen.com/content/books/principles_and_practice.pdf
- Lee, J. (2022). A meta-analysis of the effects of storytelling on English language learning. *Modern English Education*, 23(2), 33–51.
<https://doi.org/10.18095/meeso.2022.23.2.33>

- Lim, N. Z. L., Zakaria, A., & Aryadoust, V. (2022). A systematic review of digital storytelling in language learning in adolescents and adults. *Education and Information Technologies*, 27(5), 6125–6155. <https://doi.org/10.1007/s10639-021-10861-0>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press. <https://doi.org/10.1016/B978-012589042-7/50015-3>
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Papert, S., & Harel, I. (1991). Situating constructionism. In I. Harel & S. Papert (Eds.), *Constructionism* (pp. 1–11). Ablex Publishing. <https://dailypapert.com/situating-constructionism/>
- Perignat, E., & Katz-Buonincontro, J. (2019). STEAM in practice and research: An integrative literature review. *Thinking Skills and Creativity*, 31, 31–43. <https://doi.org/10.1016/j.tsc.2018.10.002>
- Pérez Agustín, M., & de-la-Peña Álvarez, C. (2025). A systematic review of the positive effects of storytelling for the linguistic competence of English as a foreign language. *Bordón, Revista de Pedagogía*, 77(1), 103–119. <https://doi.org/10.13042/Bordon.2025.106007>
- Richards, J. C. (2006). *Communicative language teaching today*. Cambridge University Press. https://www.professorjackrichards.com/wp-content/uploads/Richards-Communicative-Language.pdf?utm_source
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

GUIDED INQUIRY IN PRIMARY SCIENCE LESSONS AND THE CONCURRENT DEVELOPMENT OF SCIENCE PROCESS SKILLS, 21ST CENTURY SKILLS, AND COMPUTATIONAL THINKING: A SYSTEMATIC LITERATURE REVIEW (2016–2026)

Tracy Mensan*
Corrienna Abdul Talib

*Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia,
Johor Bahru, Malaysia
<tracy81@graduate.utm.my>*

ABSTRACT

Guided inquiry positions primary learners as active constructors of knowledge, yet evidence on its simultaneous impact on science process skills (SPS), 21st century skills (21CS) and computational thinking (CT) — competencies that together span the science, technology and mathematics strands of STEM — remains fragmented. This systematic review followed the PRISMA 2020 guideline. Six databases, including the Malaysian Citation Index and the Indonesian Garuda portal, were searched with a structured Boolean string for the period 2016–2026. After dual-reviewer screening (Cohen's $\kappa = 0.82$) and ROBINS-I appraisal, 34 empirical studies met the inclusion criteria. Because of substantial heterogeneity in instruments, designs and contexts, evidence was integrated narratively and effect sizes are reported descriptively rather than pooled: the descriptive Hedges' g ranged from moderate for CT ($g \approx 0.57$) too large for SPS ($g \approx 0.96$), with 21CS intermediate ($g \approx 0.74$). Technology-integrated guided inquiry and the 5E instructional model were the most consistently effective delivery frameworks, and intervention duration, teacher preparedness and resource availability were the key moderators. Pronounced implementation gaps appeared in rural and under-resourced ASEAN settings. The evidence base is predominantly quasi-experimental, with serious confounding risk in roughly one-quarter of studies and a preponderance of positive results consistent with possible publication bias, so the magnitudes should be read as indicative rather than definitive. The review recommends sustained teacher professional development, formative inquiry assessment embedded in existing curricula, and pre-registered longitudinal trials. For STEM education specifically, the findings suggest that guided inquiry can serve as an integrating pedagogy that develops scientific, technological and mathematical reasoning together in the primary years.

Keywords: Guided Inquiry, Computational Thinking, STEM Integration

1. Introduction

The global transformation of education in the 21st century has placed renewed emphasis on pedagogies that cultivate higher-order thinking, collaboration, and digital literacy from the earliest stages of schooling. Guided inquiry — a structured yet student-centred approach in which the teacher supplies a research question or problem and learners investigate it with scaffolded support — is widely recognised as one of the most effective pedagogies for developing these capacities in primary science education (Bybee, 2009; Pedaste et al., 2015). Large-scale syntheses have consistently associated inquiry-oriented instruction with improved science learning, although the magnitude of the benefit varies with how inquiry is enacted and with the education system in which it is embedded (Aditomo & Klieme, 2020; Minner et al., 2010). Unlike open inquiry or didactic instruction, guided inquiry occupies a productive middle

ground: it preserves learner agency while providing the scaffolding needed to prevent unproductive struggle (De Jong et al., 2024).

This review examines three outcome domains that are individually important and, for a STEM audience, jointly significant: science process skills (SPS), 21st Century Skills (21CS), and Computational Thinking (CT). These domains are not confined to a single discipline. SPS are the procedural core of the Science (S) strand; CT draws directly on the Technology (T) and Mathematics (M) strands through decomposition, algorithmic reasoning, and abstraction; and 21CS are cross-cutting competencies that a well-designed STEM lesson is expected to develop alongside content. Guided inquiry is therefore of interest not only as a science pedagogy but as a candidate vehicle for interdisciplinary STEM integration in the primary years.

Despite the proliferation of studies on inquiry-based science education, three gaps remain. First, most reviews examine a single outcome domain — typically SPS or academic achievement — without considering the simultaneous cultivation of 21CS or CT (see, for example, Minner et al., 2010; Urdanivia Alarcón et al., 2023). Second, the evidence is drawn largely from secondary and tertiary contexts, with primary populations under-represented. Third, evidence from Southeast Asian and ASEAN contexts — where cultural norms, language policy, and resource inequity shape implementation — has rarely been synthesised alongside the global literature. This review addresses these gaps by synthesising empirical evidence (2016–2026) on the effects of guided inquiry in primary science on SPS, 21CS, and CT, and by characterising the methodological quality of that evidence.

1.1 Research Questions

- a) Research Question 1: What is the overall effect of guided inquiry in primary science education on science process skills?
- b) Research Question 2: To what extent does guided inquiry foster 21st century skills (critical thinking, collaboration, communication, creativity) among primary science learners?
- c) Research Question 3: Does guided inquiry in primary science demonstrate measurable effects on computational thinking, and which instructional frameworks mediate this relationship?

2. Theoretical Background

The conceptual architecture of this review rests on three intersecting traditions: constructivism, Inquiry-Based Science Education (IBSE) theory, and the computational thinking framework.

2.1 Constructivism and Guided Inquiry

Guided inquiry is rooted in constructivist epistemology, most notably the work of Dewey (1938), Vygotsky (1978), and Bruner (1961). Dewey's insight that learning occurs through experience and reflection established inquiry as educationally legitimate. Vygotsky's zone of proximal development explains why the 'guided' dimension is indispensable: scaffolded support from a more capable guide enables learners to operate at a level they could not independently sustain. Bruner's spiral curriculum reinforces this, suggesting that learners can engage with complex ideas at progressively deeper levels when adequately supported. In the widely used taxonomy of classroom inquiry, guided inquiry occupies the level between structured and open inquiry: the teacher poses the research question and supplies the materials or procedural framework, while learners take responsibility for designing the investigation, gathering and organising data, and constructing evidence-based explanations

(Banchi & Bell, 2008). It is precisely this distribution of cognitive responsibility — a bounded problem space paired with genuine investigative agency — that makes guided inquiry well suited to primary learners, who need enough structure to avoid unproductive struggle yet still exercise the reasoning, prediction, and sense-making that drive conceptual change (Harlen, 2014; Pedaste et al., 2015). Within science education these principles were operationalised through Bybee's (2009) BSCS 5E instructional model — Engage, Explore, Explain, Elaborate, Evaluate — which was the most frequent framework in the present corpus, appearing in 23 of the 34 included studies. Recent scholarship has further refined this position by arguing that the effectiveness of inquiry-based science learning depends less on a simple contrast between inquiry and direct instruction, and more on the quality, timing, and adaptiveness of teacher guidance during investigation (de Jong et al., 2024). Areepattamannil's (2025) mini review similarly emphasises that guided inquiry must carefully orchestrate three elements — guidance, assessment, and student agency — so that learners are not merely active, but are supported toward durable conceptual understanding and inclusive participation. Emerging empirical work also continues to support the integration of guided inquiry with the 5E model, with recent studies reporting gains in science process skills, engagement, collaboration, motivation, and real-world sense-making when inquiry is structured through sequential phases of exploration, explanation, elaboration, and evaluation (Negerie et al., 2026; Siew & Chai, 2024).

2.2 Science Process Skills

Science process skills were systematically described by Harlen (2014) and elaborated by Pedaste et al. (2015) as the procedural competencies that constitute scientific practice: observing, classifying, measuring, inferring, predicting, communicating, formulating hypotheses, planning investigations, controlling variables, and interpreting data. Padilla (1990) established the still-standard distinction between basic SPS — observing, inferring, measuring, communicating, classifying, and predicting — and integrated SPS, which include controlling variables, defining operationally, formulating hypotheses, interpreting data, experimenting, and formulating models. Guided inquiry aligns with the development of both levels, requiring learners to engage in authentic investigative cycles rather than follow pre-determined instructions.

2.3 21st Century Skills

The 21st century skills framework, articulated by the Partnership for 21st Century Learning (2019) and adopted by the OECD (2018), organises competencies into four clusters: critical thinking, communication, collaboration, and creativity (the '4Cs'). Inquiry-based learning is positioned as a pedagogy that instantiates all four simultaneously — critical thinking through evidence evaluation, communication through scientific discourse, collaboration through group investigation, and creativity through experimental design (Chu et al., 2017). These competencies are precisely those that STEM education seeks to cultivate across disciplines.

2.4 Computational Thinking

Wing's (2006) account of computational thinking — formulating problems and their solutions so that the solutions can be carried out by an information-processing agent — established CT as a foundational competency. Subsequent operationalisations (Brennan & Resnick, 2012; Selby & Woollard, 2013) identified five core concepts: decomposition, pattern recognition, abstraction, algorithm design, and evaluation. Guided inquiry and CT share a problem-solving epistemology: both require learners to decompose phenomena, identify regularities, model essential features while discarding irrelevant detail, and devise systematic procedures. Psycharis and Kotzampasaki (2019) and Dickes et al. (2016, 2020) demonstrated this overlap in primary contexts, and the 5E sequence has been shown to support CT development in

elementary classrooms when its phases are used to structure programming and modelling tasks (Gao & Hew, 2022).

2.5 Guided Inquiry as STEM Integration in Primary Science

Read together, the three outcome domains map onto the STEM strands and motivate treating guided inquiry as an integrative rather than a single-subject pedagogy. The science strand is engaged directly through SPS; the technology and mathematics strands are engaged through CT, which formalises the algorithmic and quantitative reasoning implicit in careful investigation; and the engineering strand is engaged whenever inquiry is framed as iterative design and problem-solving. The 5E sequence provides a shared structure in which a single lesson can develop scientific, computational, and mathematical reasoning concurrently — the Explore phase supporting SPS and decomposition, and the Explain and Elaborate phases supporting argumentation, modelling, and transfer. This integrative reading is the lens through which the present review interprets the evidence, consistent with the mission of STEM education to connect two or more disciplinary strands around authentic problems.

2.6 Research Gaps Addressed

Three intersecting gaps motivated this review. First, no existing systematic review has synthesised evidence on all three outcome domains simultaneously within primary science, preventing an understanding of whether guided inquiry develops SPS, 21CS, and CT in mutually reinforcing ways. Second, the theoretical literature implies strong alignment between guided inquiry and CT, yet the empirical base is small ($k = 14$ studies) and heterogeneous. Third, contextual moderators — teacher capacity, language medium, resource availability, and cultural disposition toward learner questioning — have been theorised (Mahmud et al., 2018) but not systematically mapped. This review addresses all three.

3. Methodology

This review followed the PRISMA 2020 guideline (Page et al., 2021) and used the ROBINS-I framework (Sterne et al., 2016) for quality appraisal of non-randomised studies. A written protocol specifying the research questions, search strategy, eligibility criteria, extraction fields, and appraisal approach was agreed before screening. The protocol was not pre-registered on a public platform such as PROSPERO; this is acknowledged as a limitation, and pre-registration is recommended for future iterations.

3.1 Search Strategy

Six databases were searched in February 2025: Scopus, ERIC, Web of Science, the Malaysian Citation Index (MyCite), the Indonesian Garuda portal, and Google Scholar. The two regional repositories were included deliberately to surface ASEAN evidence under-indexed in the major databases. Because the search was executed in February 2025, coverage of work published in late 2025 and 2026 is necessarily partial. The following Boolean string was applied across title, abstract, and keywords:

(“guided inquiry” OR “inquiry-based learning” OR “inquiry-based science”) AND (“science process skills” OR “21st century skills” OR “computational thinking”) AND (“primary school” OR “elementary school” OR “primary education”) AND (2016:2026)

Hand-searching of the reference lists of key systematic reviews and forward and backward citation snowballing added further studies. No language restriction was applied during

identification; non-English studies were retained when an English abstract permitted eligibility assessment.

3.2 Inclusion and Exclusion Criteria

Eligibility was defined using a PICOS frame. The population was learners in primary or elementary education (approximate ages 6–12). The intervention was guided inquiry as the primary, clearly delineated instructional approach, distinguished from open inquiry and from didactic instruction. No formal comparator was required, but designs with a comparison or pre-test baseline were prioritised for effect-size estimation. The outcomes were at least one of SPS, 21CS, or CT, measured empirically. Eligible study designs were quantitative, mixed-methods, or qualitative studies published between 2016 and 2026 in a peer-reviewed journal or conference proceeding. Studies were excluded if they addressed only secondary or tertiary learners; were purely theoretical or descriptive; examined general inquiry without isolating the guided level; or were conference abstracts lacking detail for appraisal.

3.3 Study Selection

Selection proceeded in two stages. Before formal screening, the two reviewers independently coded a pilot sample of 50 records and compared decisions to calibrate the eligibility rules. They then screened all titles and abstracts independently, achieving substantial inter-rater agreement (Cohen's $\kappa = 0.82$; Landis & Koch, 1977). Disagreements were resolved through discussion, with a third reviewer arbitrating where necessary. Full texts were assessed independently against the same criteria. Of 412 records sought for retrieval, 28 could not be obtained; 384 full texts were assessed, 89 were eligible after full-text review, and 34 remained after quality appraisal. Each decision stage is documented in Figure 1.

3.4 Quality Appraisal

All 34 included studies were appraised using ROBINS-I (Sterne et al., 2016) across seven domains: confounding (D1), selection of participants (D2), classification of interventions (D3), deviations from intended interventions (D4), missing data (D5), measurement of outcomes (D6), and selective reporting (D7). The three randomised controlled trials in the CT subset were additionally appraised using the Cochrane Risk of Bias 2.0 (RoB 2) tool. Two reviewers rated each study independently and reconciled their judgements; studies at critical overall risk were excluded from the quantitative summary but retained for narrative reference.

3.5 Data Extraction and Synthesis

Two reviewers extracted data into a standardised template capturing study design, country and setting, sample size and grade level, guided-inquiry framework, outcome domain(s), assessment instrument(s), intervention duration, and reported findings including effect sizes and the statistics needed to compute them. Where a study reported several outcomes, each was recorded separately so that a single study could contribute to more than one domain. The standardised mean difference, expressed as Hedges' g , was the common effect-size metric, computed as the difference between the intervention and comparison post-test means divided by the pooled standard deviation and multiplied by the small-sample correction J , as shown in Equation 1.

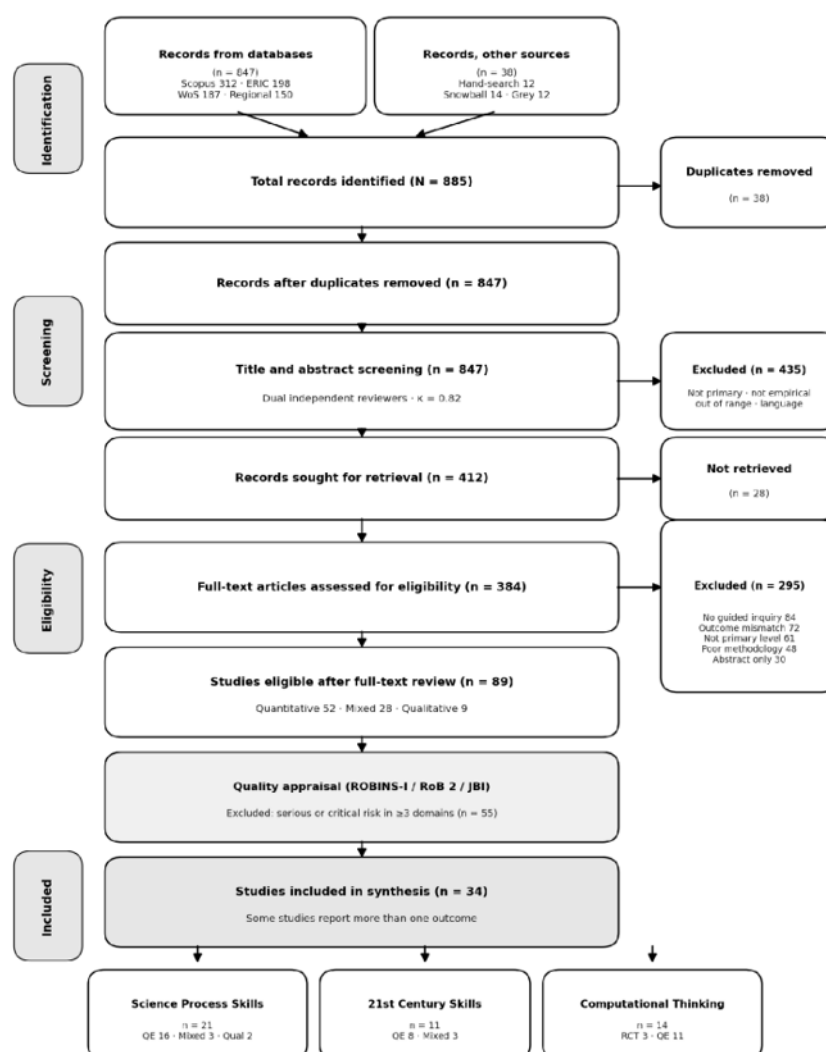
$$g = [(M_1 - M_2) / SD_{pooled}] \times J, \text{ where } J = 1 - 3 / [4(n_1 + n_2) - 9] \quad (1)$$

Where primary studies reported only t , F , or partial eta-squared, g was recovered using standard conversion formulae (Borenstein et al., 2009; Lipsey & Wilson, 2001). Magnitudes

are interpreted against Cohen's (1988) conventional benchmarks — roughly 0.2 small, 0.5 medium, and 0.8 large — while recognising that these are heuristics rather than fixed thresholds. Because of substantial methodological heterogeneity across studies, narrative synthesis was the primary mode of integration. Within each outcome domain a descriptive inverse-variance weighted mean of g is reported to summarise central tendency. This quantity is a descriptive summary, not a formal random-effects meta-analytic estimate, because pooling across such heterogeneous designs would overstate precision and several studies lacked the statistics required for variance estimation. For the same reason, formal tests of small-study effects were not computed, and the possibility of publication bias is discussed qualitatively in Section 5.3.

Figure 1

PRISMA 2020 Flow Diagram of the Study Selection Process



RCT = randomised controlled trial · QE = quasi-experimental · κ = Cohen's kappa inter-rater reliability
Adapted from the PRISMA 2020 statement (Page et al., 2022)

Note. RCT = randomised controlled trial; QE = quasi-experimental; κ = Cohen's kappa inter-rater reliability. Adapted from the PRISMA 2020 statement (Page et al., 2021).

4. Results

4.1 Study Characteristics

The 34 included studies were conducted across 14 countries, with the largest contributions from Indonesia ($k = 8$), Malaysia ($k = 6$), the Philippines ($k = 5$), China ($k = 4$), and Turkey ($k = 3$). Sample sizes ranged from 28 to 364 participants, with a median n of 78 — a figure that signals how small many of the primary studies were. Quasi-experimental designs dominated ($k = 28$), followed by mixed-methods ($k = 3$) and qualitative ($k = 3$) studies; the three randomised trials were confined to the CT cluster. The 5E model was the most frequent framework ($k = 23$), ahead of POGIL ($k = 5$) and researcher-developed protocols ($k = 6$). Intervention length ran from two weeks to a full semester, with six weeks the modal duration. Table 1 summarises these characteristics.

Table 1

Characteristics of Included Studies by Outcome Domain

Outcome Domain	Studies (k)	Predominant Design	Median n	Most Common Framework
Science process skills	21	Quasi-experimental (18)	82	5E model (Indonesia)
21st century skills	11	Quasi-experimental (8)	76	5E model (Philippines)
Computational thinking	14	Quasi-experimental (11)	65	5E model / technology-enhanced
Multiple outcomes	12	Mixed methods (6)	94	5E model (Malaysia)

Note. Columns sum to more than 34 because 12 studies reported more than one outcome domain. Compiled by the authors.

4.2 Effects on Science Process Skills (RQ1)

Twenty-one studies, covering 1,742 primary learners, reported SPS outcomes. Eighteen reported positive effects, with effect sizes spanning a small $g \approx 0.31$ to a large $g \approx 1.42$; the descriptive inverse-variance weighted mean was $g \approx 0.96$ (95% CI [0.78, 1.14]). The most consistently improved sub-skills — observing, predicting, and communicating findings — align with the Explore and Elaborate phases of the 5E cycle. Integrated skills such as hypothesis formulation, experimental design, and data interpretation improved more modestly and were more sensitive to duration: interventions of six weeks or longer produced larger gains than shorter ones (Kamarudin et al., 2024; Sandika & Fitrihidajati, 2018), although the duration signal draws partly on studies with older learners and should therefore be treated as suggestive. Indonesian module-development studies reported comparable gains in basic SPS, albeit at high risk of bias (Astuti et al., 2018). Context moderated these effects: rural Malaysian and Indonesian settings reported smaller effects than urban comparators, a pattern attributed to material resources and teacher readiness (Halim et al., 2018; Mahmud et al., 2018), and similar resource-related attenuation has been documented in other under-resourced systems (Chengere et al., 2025). Low-cost, curriculum-aligned laboratory kits and e-modules partly compensated for these constraints (Wong & Sim, 2022), although digital provision introduced new dependencies on connectivity and device access in interior Borneo schools.

4.3 Effects on 21st Century Skills (RQ2)

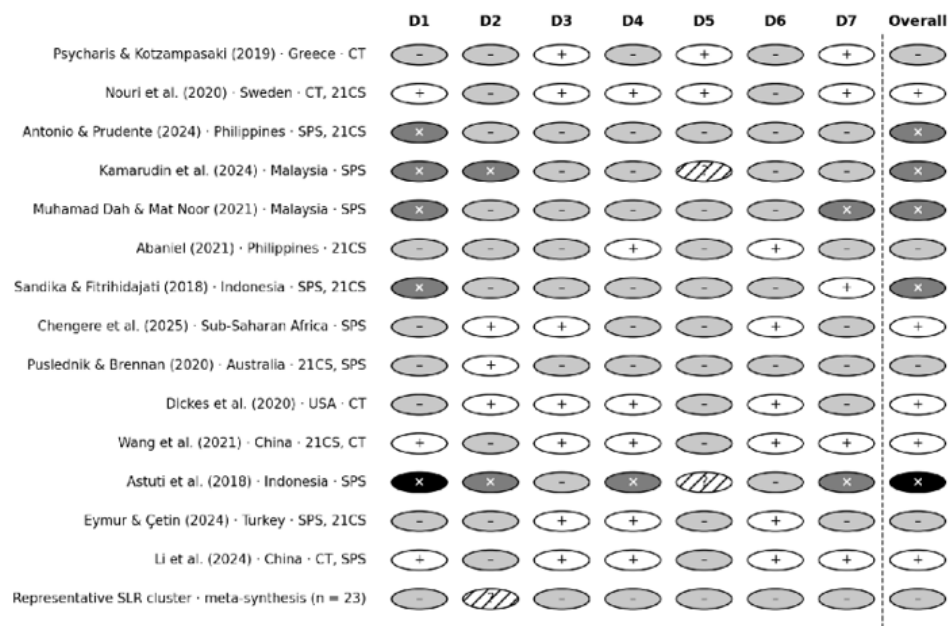
Eleven studies reported 21CS outcomes, with the most frequently measured sub-skills being critical thinking ($k = 9$), collaboration ($k = 8$), and communication ($k = 6$); creativity received the least attention ($k = 3$). The descriptive weighted mean was $g \approx 0.74$ (95% CI [0.58, 0.90]). Benefits for critical thinking were most pronounced when the Explain and Elaborate phases were well implemented — phases that require learners to publicly construct and defend explanations (Eymur & Çetin, 2024). Collaboration and communication gains were robust across settings, although the clearest evidence comes from open-inquiry implementations with older learners (Abaniel, 2021; Puslednik & Brennan, 2020), which limits direct transfer to guided inquiry in the primary years. These patterns are broadly consistent with meta-analytic evidence that inquiry-based approaches raise higher-order thinking outcomes in science (Antonio & Prudente, 2024). However, several studies noted that cultural norms — deference to teacher authority and social-harmony dispositions in Southeast Asian classrooms (Halim et al., 2018) — initially suppressed spontaneous questioning, requiring explicit facilitation strategies to unlock these gains. Structured questioning routines are one such strategy: a Malaysian teacher-researcher study found that an explicit question-formulation technique enabled pupils to generate investigable questions that they had not produced spontaneously (Muhamad Dah & Mat Noor, 2021).

4.4 Effects on Computational Thinking (RQ3)

Fourteen studies examined CT outcomes — the smallest cluster and the one with the most methodological heterogeneity. CT was measured with diverse instruments, including the Computational Thinking Test, Bebras tasks, researcher-developed rubrics, and portfolio assessment, a fragmentation that limits comparability. Despite this, a positive directional trend was evident in 12 of the 14 studies, with a descriptive weighted mean $g \approx 0.57$ (95% CI [0.38, 0.76]), a moderate effect. Technology integration substantially mediated CT: studies embedding guided inquiry within visual programming environments, agent-based modelling platforms, or robotics reported larger effects than ‘unplugged’ guided inquiry alone (Dickes et al., 2020; Psycharis & Kotzampasaki, 2019). The 5E model was particularly associated with CT development when the Explore phase incorporated model-building or algorithm-driven tasks (Gao & Hew, 2022) — an illustration of the technology and mathematics strands being engaged through a science lesson. Studies that treated CT and 21CS as co-developing rather than separate outcomes reported the same pattern (Nouri et al., 2020).

4.5 Risk of Bias

Across the 34 studies, confounding (D1) was the most prevalent bias domain: 22 studies used intact class groups without randomisation, making teacher effect, prior ability, and school resource levels unavoidable confounders. Selective reporting (D7) was serious in eight studies. Overall risk judgements were low ($k = 6$, 17.6%), moderate ($k = 16$, 47.1%), serious ($k = 10$, 29.4%), and critical ($k = 2$, 5.9%); the two critical studies were retained for narrative reference only. Figure 2 presents the domain-level appraisal for a representative set of studies, and Table 2 the overall distribution.

Figure 2**ROBINS-I Risk-of-Bias Assessment for 15 Representative Studies**

D1 confounding - D2 selection of participants - D3 classification of interventions - D4 deviations from intended interventions
 D5 missing data - D6 measurement of outcomes - D7 selective reporting. ROBINS-I tool (Sterne et al., 2016).
 Note: 15 representative studies are shown; the remaining 19 included studies were appraised under the same protocol.

○ Low (+) ○ Moderate (-) ● Serious (x) ● Critical (x) ○ No information (?)

Note. Appraisal conducted by the authors using ROBINS-I (Sterne et al., 2016). Fifteen representative studies are shown; the remaining 19 included studies were appraised under the same protocol. SPS = science process skills; 21CS = 21st century skills; CT = computational thinking.

Table 2**ROBINS-I Overall Risk-of-Bias Summary Across the 34 Included Studies**

Risk Level	Studies (k)	%	Primary Bias Domain	Use in Synthesis
Low	6	17.6	Minor selection bias (D2)	Full synthesis
Moderate	16	47.1	Confounding (D1)	Full synthesis
Serious	10	29.4	Confounding and reporting (D1, D7)	Narrative only
Critical	2	5.9	Confounding and missing data (D1, D5)	Excluded

Note. Percentages are rounded and sum to 100. Appraisal conducted by the authors using ROBINS-I (Sterne et al., 2016). D1 = confounding; D2 = selection of participants; D5 = missing data; D7 = selective reporting.

5. Discussion

The synthesis indicates that guided inquiry in primary science is associated with positive effects across all three outcome domains, with the largest descriptive effect for SPS ($g \approx 0.96$) and the smallest for CT ($g \approx 0.57$). This gradient is consistent with the theoretical framework: SPS is the most proximal outcome to guided inquiry, being intrinsic to the inquiry process, whereas CT is a more distal transfer outcome that requires deliberate instructional design (Twizeyimana et al., 2024). The gradient may also reflect the greater maturity of the SPS evidence base relative to the small, instrument-fragmented CT literature. These magnitudes should be read cautiously rather than celebrated: they are descriptive summaries drawn from predominantly quasi-experimental studies, and the gap between the largest and smallest effects is well within the range that design and measurement differences alone could produce.

5.1 Implementation Conditions and Moderators

Three implementation conditions recurred. First, intervention duration matters: studies of six weeks or more systematically outperformed shorter interventions, suggesting that inquiry dispositions need time to consolidate — a finding with direct implications for the compressed weekly science allocations in Malaysian primary schools (Ministry of Education Malaysia, 2013). Second, technology integration was a positive moderator for CT but a double-edged one in under-resourced contexts where device access and connectivity are unreliable. Third, teacher preparedness mediated all three outcomes: studies in which teachers received inquiry coaching or belonged to professional learning communities reported consistently larger effects, echoing wider evidence that the quality of enacted inquiry depends heavily on teacher preparation and instructional decision-making (Cairns, 2019; Strat et al., 2024).

5.2 Cultural Moderation in ASEAN Contexts

The cultural findings are salient for ASEAN settings. The suppressive effect of deference norms on collaborative inquiry — documented in Malaysia (Halim et al., 2018) and the Philippines (Abaniel, 2021) — suggests that guided inquiry cannot be transplanted without explicit strategies to normalise learner questioning and peer challenge. Question walls, structured think-pair-share protocols, and anonymous question-posting routines emerge from the literature as viable structural responses. These observations rest on a small number of studies and should be treated as hypotheses for targeted investigation rather than established moderators.

5.3 Limitations of the Evidence Base and Publication Bias

Several features of the corpus temper its conclusions. The dominance of intact-class quasi-experiments means teacher effects, prior ability, and resourcing are frequently confounded with the intervention, and the median sample n of 78 leaves many studies underpowered for the moderate effects they report. Outcome measurement, especially for CT, is fragmented across non-comparable instruments, which inflates apparent heterogeneity. A further limitation concerns population fit: several of the most frequently cited studies in the 21CS cluster were conducted with secondary or tertiary learners, or examined open rather than guided inquiry, so their contribution to conclusions about guided inquiry in the primary years is indirect.

The distribution of results is itself informative. With 18 of 21 SPS studies and 12 of 14 CT studies reporting positive effects, and with many small samples, the corpus shows the profile in which small-study and publication effects are most likely: small underpowered studies with null or negative results tend not to be published, biasing the visible literature upward. Because the synthesis is narrative and many studies lacked the statistics for variance estimation, this could not be tested formally through funnel-plot asymmetry or related methods. The design

choices partially mitigate the risk — the search included grey literature and two regional repositories that index work less subject to international publication filters, and no language restriction was applied. Even so, the descriptive weighted means in Section 4 should be regarded as plausibly optimistic upper estimates rather than unbiased point estimates, and a future meta-analysis with formal small-study diagnostics is warranted once instrument harmonisation makes pooling defensible.

5.4 Research Gaps Identified

Three gaps follow. First, there is a near-total absence of longitudinal studies tracking CT development through guided inquiry beyond a single semester. Second, CT assessment instruments for primary inquiry are fragmented and non-comparable — 14 studies used 11 different tools — so a validated, context-sensitive CT measure is urgently needed. Third, the intersection of language medium and inquiry effectiveness, particularly relevant in Malaysian Dual Language Programme schools, has been observed but not systematically studied (Mahmud et al., 2018); teachers whose scientific vocabulary in English is limited may suppress the quality of the Explain and Elaborate phases precisely where CT and 21CS gains depend most on rich discourse. Taken together, the evidence supports a unifying reading in which guided inquiry behaves less like three separate interventions and more like a single integrative STEM pedagogy whose 5E phases each contribute to several outcomes. This reading is a hypothesis consistent with the present evidence rather than one it conclusively establishes; testing it directly would require studies designed to measure the three domains concurrently, which remain rare.

6. Conclusion

This review synthesised 34 empirical studies on the effects of guided inquiry in primary science on science process skills, 21st century skills, and computational thinking. Guided inquiry was associated with positive effects across all three domains — largest for SPS ($g \approx 0.96$), intermediate for 21CS ($g \approx 0.74$), and moderate for CT ($g \approx 0.57$) — with the 5E instructional model the most common and apparently most effective framework and technology-enhanced inquiry amplifying CT outcomes. For STEM education, the pattern suggests that a single well-designed inquiry lesson can develop scientific, technological, and mathematical reasoning together, supporting the treatment of guided inquiry as an integrative STEM pedagogy in the primary years.

These conclusions rest on an evidence base that is predominantly quasi-experimental, vulnerable to confounding, and likely skewed toward positive results, so the magnitudes should be read as indicative rather than definitive. Four limitations of the review itself should also be borne in mind: the protocol was not pre-registered; evidence was synthesised narratively rather than meta-analytically, so the weighted means are descriptive; the February 2025 search makes coverage of 2025–2026 work incomplete; and reliance on English abstracts for non-English records may have led to omissions. Future research should prioritise pre-registered, cluster-randomised trials with adequate samples, mixed-methods designs that capture mechanism alongside magnitude, and longitudinal tracking across academic years. Policymakers and curriculum developers in Malaysia and the broader ASEAN region should support guided inquiry with sustained, inquiry-specific professional development, formative assessment aligned to existing curriculum standards, and resource provision that narrows the technology gap between urban and rural learners.

Acknowledgment

The authors thank the reviewers and colleagues at Universiti Teknologi Malaysia who contributed to screening and appraisal.

References

- Abaniel, A. (2021). Enhanced conceptual understanding, 21st century skills and learning attitudes through an open inquiry learning model in physics. *Journal of Technology and Science Education*, 11(1), 30–43. <https://doi.org/10.3926/jotse.1004>
- Aditomo, A., & Klieme, E. (2020). Forms of inquiry-based science instruction and their relations with learning outcomes: Evidence from high and low-performing education systems. *International Journal of Science Education*, 42(4), 504–525. <https://doi.org/10.1080/09500693.2020.1716093>
- Antonio, R. P., & Prudente, M. S. (2024). Effectiveness of inquiry-based approaches on students' higher-order thinking skills in science: A meta-analysis. *International Journal of Education in Mathematics, Science and Technology*, 12(1), 251–281. <https://doi.org/10.46328/ijemst.3216>
- Areepattamannil, S. (2025). Guided inquiry in school science: A mini review of orchestration, assessment, and AI. *Frontiers in Education*, 10, Article 1534358. <https://doi.org/10.3389/feduc.2025.1534358>
- Astuti, M. W., Hartini, S., & Mastuang, M. (2018). Pengembangan modul IPA dengan menggunakan model pembelajaran inkuiri terbimbing pada materi suhu dan kalor untuk melatih keterampilan proses sains [Development of a science module using the guided inquiry learning model on temperature and heat to train science process skills]. *Berkala Ilmiah Pendidikan Fisika*, 6(2), 205–218. <https://doi.org/10.20527/bipf.v6i2.4934>
- Banchi, H., & Bell, R. (2008). The many levels of inquiry. *Science and Children*, 46(2), 26–29. <https://eric.ed.gov/?id=EJ815766>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. <https://doi.org/10.1002/9780470743386>
- Brennan, K., & Resnick, M. (2012, April 13–17). New frameworks for studying and assessing the development of computational thinking [Paper presentation]. Annual Meeting of the American Educational Research Association, Vancouver, BC, Canada. https://web.media.mit.edu/~kbrennan/files/Brennan_Resnick_AERA2012_CT.pdf
- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31(1), 21–32.
- Bybee, R. W. (2009). *The BSCS 5E instructional model and 21st century skills* [Paper prepared for the National Academies Board on Science Education]. National Academies Press.
- Cairns, D. (2019). Investigating the relationship between instructional practices and science achievement in an inquiry-based learning environment. *International Journal of Science Education*, 41(15), 2113–2135. <https://doi.org/10.1080/09500693.2019.1660927>
- Chengere, A. M., Bono, B. D., Zinabu, S. A., & Jilo, K. W. (2025). Enhancing secondary school students' science process skills through guided inquiry-based laboratory activities in biology. *PLOS ONE*, 20(4), Article e0320692. <https://doi.org/10.1371/journal.pone.0320692>

- Chu, S. K. W., Reynolds, R. B., Tavares, N. J., Notari, M., & Lee, C. W. Y. (2017). *21st century skills development through inquiry-based learning: From theory to practice*. Springer. <https://doi.org/10.1007/978-981-10-2481-8>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- De Jong, T., Lazonder, A. W., Chinn, C. A., Fischer, F., Gobert, J., Hmelo-Silver, C. E., Koedinger, K. R., Krajcik, J. S., Kyza, E. A., Linn, M. C., Pedaste, M., Scheiter, K., & Zacharia, Z. C. (2024). Beyond inquiry or direct instruction: Pressing issues for designing impactful science learning opportunities. *Educational Research Review*, 44, Article 100623. <https://doi.org/10.1016/j.edurev.2024.100623>
- Dewey, J. (1938). *Experience and education*. Macmillan.
- Dickes, A. C., Farris, A. V., & Sengupta, P. (2020). Sociomathematical norms for integrating coding and modeling with elementary science: A dialogical approach. *Journal of Science Education and Technology*, 29(1), 35–52. <https://doi.org/10.1007/s10956-019-09795-7>
- Dickes, A. C., Sengupta, P., Farris, A. V., & Basu, S. (2016). Development of mechanistic reasoning and multilevel explanations of ecology in third grade using agent-based models. *Science Education*, 100(4), 734–776. <https://doi.org/10.1002/sce.21217>
- Eymur, G., & Çetin, P. S. (2024). Investigating the role of an inquiry-based science lab on students' scientific literacy. *Instructional Science*, 52, 743–760. <https://doi.org/10.1007/s11251-024-09672-w>
- Gao, X., & Hew, K. F. (2022). Toward a 5E-based flipped classroom model for teaching computational thinking in elementary school: Effects on student computational thinking and problem-solving performance. *Journal of Educational Computing Research*, 60(2), 512–543. <https://doi.org/10.1177/07356331211037757>
- Halim, L., Ramli, M., & Mohamad Nasri, N. (2018). Cultural influences in inquiry-based science learning (IBSL) practices: Malaysia and Indonesia context. In Proceedings of the 1st PGSD International Conference on Education (pp. 1–8).
- Harlen, W. (2014). Helping children's development of inquiry skills. *Inquiry in Primary Science Education*, 1(1), 5–19. <https://ipsejournal.com/wp-content/uploads/2015/03/3-ipse-volume-1-no-1-wynne-harlen-p-5-19.pdf>
- Kamarudin, M. Z., Mat Noor, M. S. A., & Omar, R. (2024). A scoping review of the effects of a technology-integrated, inquiry-based approach on primary pupils' learning in science. *Research in Science & Technological Education*, 42(3), 828–847. <https://doi.org/10.1080/02635143.2022.2138847>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE.
- Mahmud, S. N. D., Mohamad Nasri, N., Samsudin, M. A., & Halim, L. (2018). Science teacher education in Malaysia: Challenges and way forward. *Asia-Pacific Science Education*, 4, Article 8. <https://doi.org/10.1186/s41029-018-0026-3>
- Ministry of Education Malaysia. (2013). *Malaysia education blueprint 2013–2025 (preschool to post-secondary education)*. Ministry of Education Malaysia.
- 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>
- Muhamad Dah, N., & Mat Noor, M. S. A. (2021). Facilitating pupils' questioning skills in open inquiry learning through an Investigable Question Formulation Technique (IQFT).

- Journal of Mathematics and Science Teacher*, 1(2), Article em005.
<https://doi.org/10.29333/mathsciteacher/11283>
- Negerie, D., Habtamu, T., & Tafesse, T. (2026). Effect of guided inquiry with the 5E model on students' engagement in biology at secondary schools in Mattu town, Southwest Ethiopia. *Discover Education*, 5, Article 203. <https://doi.org/10.1007/s44217-026-01249-8>
- Nouri, J., Zhang, L., Mannila, L., & Norén, E. (2020). Development of computational thinking, digital competence and 21st century skills when learning programming in K-9. *Education Inquiry*, 11(1), 1–17. <https://doi.org/10.1080/20004508.2019.1627844>
- OECD. (2018). *The future of education and skills: Education 2030*. OECD Publishing.
- Padilla, M. J. (1990). *The science process skills* (Research Matters — to the Science Teacher No. 9004). National Association for Research in Science Teaching. <https://narst.org/research-matters/science-process-skills>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
- Partnership for 21st Century Learning. (2019). *Framework for 21st century learning definitions*. Battelle for Kids. https://static.battelleforkids.org/documents/p21/P21_Framework_DefinitionsBFK.pdf
- Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14, 47–61. <https://doi.org/10.1016/j.edurev.2015.02.003>
- Psycharis, S., & Kotzampasaki, E. (2019). The impact of a STEM inquiry game learning scenario on computational thinking and computer self-confidence. *EURASIA Journal of Mathematics, Science and Technology Education*, 15(4), Article em1689. <https://doi.org/10.29333/ejmste/103071>
- Puslednik, L., & Brennan, M. (2020). Enabling students to engage in scientific inquiry through an open inquiry approach enhances 21st-century skills. *Teaching Science*, 66(2), 18–29.
- Sandika, B., & Fitrihidajati, H. (2018). Improving creative thinking skills and scientific attitude through inquiry-based learning in basic biology lecture toward students of biology education. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, 4(1), 23–28. <https://doi.org/10.22219/jpbi.v4i1.5326>
- Selby, C. C., & Woollard, J. (2013). *Computational thinking: The developing definition*. University of Southampton. <https://eprints.soton.ac.uk/356481/>
- Siew, N. M., & Chai, W. L. (2024). The integration of 5E inquiry-based learning and group investigation model: Its effects on level four science process skills of Form Four students. *Problems of Education in the 21st Century*, 82(1), 133–148. <https://doi.org/10.33225/pec/24.82.133>
- Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., ... Higgins, J. P. T. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, 355, Article i4919. <https://doi.org/10.1136/bmj.i4919>

- Strat, T. T. S., Henriksen, E. K., & Jegstad, K. M. (2024). Inquiry-based science education in science teacher education: A systematic review. *Studies in Science Education*, 60(2), 191–249. <https://doi.org/10.1080/03057267.2023.2207148>
- Twizeyimana, E., Shyiramunda, T., Dufitumukiza, B., & Niyitegeka, G. (2024). Teaching and learning science as inquiry: An outlook of teachers in science education. *SN Social Sciences*, 4, Article 40. <https://doi.org/10.1007/s43545-024-00846-4>
- Urdanivia Alarcón, D. A., Talavera-Mendoza, F., Rucano Paucar, F. H., Cayani Caceres, K. S., & Machaca Viza, R. (2023). Science and inquiry-based teaching and learning: A systematic review. *Frontiers in Education*, 8, Article 1170487. <https://doi.org/10.3389/feduc.2023.1170487>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. <https://doi.org/10.1145/1118178.1118215>
- Wong, H. T., & Sim, S. F. (2022). A curriculum-based laboratory kit for flexible teaching and learning of practical chemistry. *Chemistry Teacher International*, 4(4), 343–353. <https://doi.org/10.1515/cti-2022-0014>

FACULTY ATTITUDES TOWARDS OUTCOME-BASED EDUCATION IMPLEMENTATION IN LAO PUBLIC UNIVERSITIES: A CASE STUDY OF ENGINEERING EDUCATION

Ayyia Machang^{1,3*}
Yongxia Ma¹
Houmlack Mingboubpha²

¹*School of Education, Beijing Institute of Technology, PR China*

²*School of Economics, Beijing Institute of Technology, PR China*

³*Department of Higher Education, Ministry of Education and Sports, Lao PDR*
<ayyiachang@gmail.com>

ABSTRACT

The development of higher education curricula aligned with regional and international frameworks to facilitate student mobility, credit transfer, and development of new graduates with relevant knowledge and skills to meet labour market demands has become a crucial transition from traditional education to Outcomes-Based Education (OBE) in Laos. The current study investigates faculty members' attitudes towards OBE implementation, identifies the highest and lowest attitude components, and examines the effects of teaching experience and OBE training on engineering faculty members' attitudes at public universities in Laos. The study employs a quantitative approach using a structured questionnaire administered to 267 faculty members. The findings show that 77% of instructors agree or strongly agree with the OBE approach. Beliefs, feelings, readiness and acceptance are generally positive, exceeding a 70% agreement level. However, knowledge of the OBE approach remains low relative to this threshold, with a combined agreement level of 64.32%. The mean score for overall attitude is 3.98, with acceptance being the highest ($A_c = 4.13$) and knowledge ($K_n = 3.75$) the lowest. The findings also show that teaching experience and OBE training exhibit a statistically significant effect on faculty members' attitudes towards OBE implementation in engineering education at Lao public universities.

Keywords: Outcomes-Based Education (OBE), Faculty Members' Attitudes, Engineering Education, Lao Public University.

1. Introduction

The development of higher education curricula aligned with regional and international frameworks to facilitate student mobility, credit transfer, and the development of new graduates with relevant knowledge and skills to meet labour market demands has become a crucial transition from traditional education to Outcomes-Based Education (OBE) in higher education in Laos. A survey study on higher education graduates' employment in Laos found that 71.61% of employers believed that the graduates' knowledge and skills were relevant to their jobs, while 27.5% considered them only partly relevant and 0.83% reported no relevance (Duangdala et al., 2023). The main concerns raised by employers include a lack of practical skills, insufficient knowledge, weak alignment with working conditions, and overly general curricula. These findings indicate that the traditional education approaches remain insufficiently aligned with labour market demands (Ministry of Education and Sports [MOES], 2011). Consequently, new graduates are inadequately prepared to meet the requirements of the modern engineering sector. This issue raises concerns about the capacity of the higher education system to support socio-economic development and global competitiveness. The

need to address these challenges has led to calls for reform in engineering education to equip new graduates with the competencies required for innovation and problem-solving in both domestic and international contexts (Lavi & Bagiati, 2022). Outcomes-Based Education (OBE) has emerged as a widely adopted approach that emphasises clearly defined learning outcomes and student-centred learning processes (Pradhan, 2021).

Higher education curricula have been required to align with the four domains specified in the National Curriculum Standards since Laos' MOES adopted them in 2020. These domains include knowledge, skills, practical skills, and social skills (Ministry of Education and Sports [MOES], 2020). The Lao National Qualification Framework also requires all HEIs to specify learning outcomes to enhance teaching quality and alignment with labour market demands (Government of the Lao People's Democratic Republic [Government of Lao PDR], 2021). Approximately 20 higher education institutions and 300 undergraduate programmes have transitioned to the OBE approach over the past five years. However, OBE remains relatively new in practice, and its implementation at the faculty level is still uncertain and uneven, particularly in engineering programmes. Moreover, the existing literature provides limited empirical evidence on OBE implementation in engineering education within the Lao context; there is a lack of systematic analysis of faculty attitudes, which play a critical role in curriculum reform. Although some programmes in Laos have been assessed under AUN-QA standards, accreditation remains limited, and evidence on how faculty members perceive and implement OBE is still insufficient (ASEAN University Network [AUN], 2022). Thus, this study aims to address this gap by examining faculty members' attitudes towards OBE implementation in engineering education at public universities in Laos.

- a) Research Question 1: What are faculty members' attitudes towards OBE implementation in engineering education?
- b) Research Question 2: Which attitude components receive the highest and lowest levels of agreement among faculty members?
- c) Research Question 3: Do teaching experience and OBE training affect faculty members' attitudes towards OBE implementation in engineering education?

2. Literature Review

2.1 Concepts of Outcomes-Based Education

Outcome-Based Education (OBE) shifts the educational emphasis from instructional inputs to what learners are expected to accomplish. While more recent research emphasises predefined and measurable outcomes (Allo et al., 2024). Early studies characterise OBE in terms of clearly defined learning outcomes and student learning (Spady, 1994; Giessen-Hood, 1999; Davis, 2003). Collectively, these perspectives position learning outcomes as the starting point for curriculum design. OBE also integrates student-centred instruction, assessment, and competence development (Akramy, 2021; Yusoff et al., 2014; Dano, 2024). Thus, rather than considering teaching and assessment independently, OBE aligns intended outcomes with instructional and assessment practices (Khanna & Mehrotra, 2019; Yusof et al., 2017; Pradhan, 2021). Such alignment may increase students' knowledge and willingness to study (Sunra et al., 2024).

2.2 OBE Curriculum Versus Traditional Curriculum: Shifting Paradigm

The emphasis of the learning process is the primary difference between OBE and conventional education. Traditional curricula prioritise instructional inputs and information delivery, while

OBE emphasises the skills students can exhibit (An, 2014; Sawant, 2016). This transition is consistent with constructive alignment and backward curriculum design, which guide teaching and assessment based on learning outcomes (Biggs, 1996; Harden et al., 1999). Although OBE fosters student engagement and motivation (Deepak & Venishri, 2018), its implementation requires faculty members to adapt traditional teaching and assessment practices. Therefore, the move to OBE involves both a curricular reform and a change in instructors' pedagogical roles.

2.3 Outcome-Based Education in Engineering Programmes

OBE remains particularly relevant to engineering education because engineering curricula must constantly evolve in response to industrial and technological advancements (Magenuka & Sibanda, 2023). To enable successful implementation, assessment practices should align with clearly defined learning outcomes (Jadhav et al., 2020). Implementation can still be uneven and may improve over time (Canini & Dalis, 2016). However, research on OBE's effect on engineering education has identified positive impacts on student academic performance, technical skills, and transferable skills (Borsoto et al., 2014; Hegde, 2014; Isa et al., 2017; Macatangay et al., 2016). External stakeholder feedback and regular reviews also help maintain outcome alignment (Ali, 2023; Puerto- López et al., 2023). While these findings show the potential in OBE, they also suggest that effective implementation depends on faculty members' ability to translate learning outcomes into teaching and assessment practices.

2.4 Faculty Members' Attitude Towards OBE Approach

Previous literature shows that faculty members' reactions to OBE are mostly positive. Studies from Alimyar (2020); Baguio (2019); and Ortega and Ortega-Dela Cruz (2016) surveyed faculty members and found that they were happy, willing to apply OBE, and confident that OBE could meet students' learning needs and demands of the labour market. Positive attitudes do not always translate into sufficient knowledge to encourage effective implementation. While Jansen (1998) emphasises the conceptual and terminological challenges teachers encounter, Katawazai (2021) finds strong conviction in OBE combined with insufficient awareness of its practical implementation. The research consequently exposes a significant tension: faculty may endorse OBE while lacking the skills necessary to apply it successfully. Examining attitudes across knowledge, beliefs, emotions, readiness, and acceptance rather than using a single measure is supported by this difference.

2.5 Faculty Members' Personnel Factor Affect OBE Implementation

Factors at the institutional and faculty levels may also affect the implementation of OBE. Resistance to curriculum change can stem from heavy workloads and a lack of knowledge (Cheung & Man Wong, 2012), while limited awareness and negative attitudes can create additional barriers to implementation (Akhmadeeva et al., 2013). Specifically, inadequate training can lead to misalignment among learning objectives, instruction, and assessment being misaligned (Mogashoa, 2013). Institutional constraints and resistance to learner-centred teaching may exacerbate these issues (Spady, 1994; Yasmin & Yasmeen, 2021).

2.6 Literature Gaps

While evidence suggests that OBE is generally endorsed by faculty members, existing studies do not address such endorsement related to knowledge and ability to implement OBE. Research typically isolates variables such as implementation challenges, perceived effectiveness, or overall attitudes. It remains unclear whether faculty acceptance and readiness exists alongside a lack of conceptual knowledge of OBE. Additionally, whether professional experience (including years of teaching and OBE training) is associated with

systematic differences among these attitudinal measures remains unknown. This gap limits the ability of higher education institutions transitioning from content-based curricula to understand their faculty members' attitudes towards OBE. Since OBE was recently introduced to Laos' national curriculum and accreditation standards, empirical data on faculty perceptions remain limited, particularly within engineering education. To address this gap, this study investigates faculty attitudes towards OBE, including knowledge, beliefs, feelings, readiness, and acceptance. Additionally, this study examines whether years of teaching experience and OBE training correspond to systematic differences in these factors.

3. Theoretical Framework

This research conceptualises faculty attitudes towards OBE across five interrelated dimensions: knowledge, beliefs, feelings, readiness, and acceptance. Knowledge reflects an understanding of OBE concepts and practices (Harden et al., 1999), while beliefs and sentiments capture faculty members' cognitive and emotional appraisals of OBE (Baguio, 2019). Readiness and acceptance reflect their willingness to embrace and apply the approach (Ortega & Ortega-Dela Cruz, 2016).

Although these dimensions may not develop uniformly, they are interrelated. While a lack of knowledge may impede practical application, positive attitudes and sentiments may promote preparedness and acceptance. Faculty may therefore embrace OBE without fully comprehending its practical application (Jansen, 1998; Katawazai, 2021). Teaching experience and OBE training may further influence these dimensions by enhancing faculty confidence and implementation capacity (Cheung & Man Wong, 2012; Mogashoa, 2013).

Accordingly, the study compares the five attitudinal dimensions to identify relative strengths and weaknesses and evaluates whether teaching experience and OBE training are associated with variations across these dimensions. This framework establishes a direct link between the study's descriptive, comparative, and ANOVA analyses and faculty attitudes.

4. Research Methodology

The research employs a quantitative, cross-sectional survey design to examine faculty members' attitudes towards the OBE implementation in engineering programmes at public universities in Laos. The quantitative design is appropriate for assessing attitudinal dimensions in a structured manner and examining variations across faculty characteristics. The study analyses survey responses collected at a single point in time, allowing for the identification of systematic patterns in attitudes and associated factors, such as teaching experience and participation in OBE-related training.

4.1 Research Instrument

The instrument used in this study consists of 26 items that capture faculty members' opinions regarding the use of OBE in HEIs. These items are categorised into five attitudinal dimensions: knowledge, beliefs, feelings, preparedness, and acceptance of the OBE approach. The questionnaire items were adapted from previous studies examining faculty attitudes towards OBE implementation (Baguio, 2019; Chowdhury & Das, 2022; Katawazai, 2021; Loreto, 2018; Ortega & Ortega-Dela Cruz, 2016). The survey uses a five-point Likert scale where respondents select responses ranging from 1 "strongly disagree" to 5 "strongly agree" for each item. Higher scores indicate more positive attitudes. The questionnaire items were reviewed prior to data collection to ensure clarity, relevance, and consistency with the OBE policy context of Lao higher education. The reliability and suitability of the instrument were further

assessed using Cronbach's alpha, the Kaiser-Meyer-Olkin (KMO) measure, and Bartlett's test of sphericity.

4.2 Sample of Research

- a) Selection of universities: Four public universities were purposely selected based on specific criteria, including the adoption of OBE, the presence of engineering faculties, and active teaching and learning in engineering programmes. This selection focused on public universities because they share similar functions, roles, and structures, and have similar experiences with engineering education in Laos.
- b) Participants: The participants were faculty members teaching engineering at the public universities. Faculty members were chosen as the target respondents because they are directly engaged in translating OBE curriculum requirements into engineering teaching and assessment practices. More than 300 questionnaires were distributed, and 267 valid responses were retained for analysis

4.3 Data Collective Procedures

The researchers initially submitted a proposal letter requesting approval from the Department of Higher Education (DHE) within the Ministry of Education and Sports, Laos. After reviewing and approving the proposed research objectives, DHE issued a letter to the presidents of the four public universities asking them to facilitate data collection. After receiving approval from each university, the researchers travelled to the universities to distribute questionnaires to faculty members. The questionnaires were distributed to faculty members in the Faculty of Engineering at the National University of Laos, the Faculty of Engineering at Souphanouvong University, the Faculty of Engineering and Faculty of Technology and Information at Savannakhet University and the Faculty of Engineering at Champasack University. Participation was entirely voluntary, and participants were informed of the study's academic purpose. Responses were used exclusively for research purposes and analysed without identifying individual participants.

4.4 Data Analysis Procedures

Data analysis and coding were conducted using IBM SPSS Statistics version 27. The analysis was conducted in multiple stages. The dataset underwent an initial evaluation to ensure data quality, including checks for inconsistent entries, unusual values, and missing data. The researchers carried out descriptive statistical analyses of faculty members' attitudes towards OBE implementation by calculating means and standard deviations, as well as computing frequencies and percentages for the five dimensions.

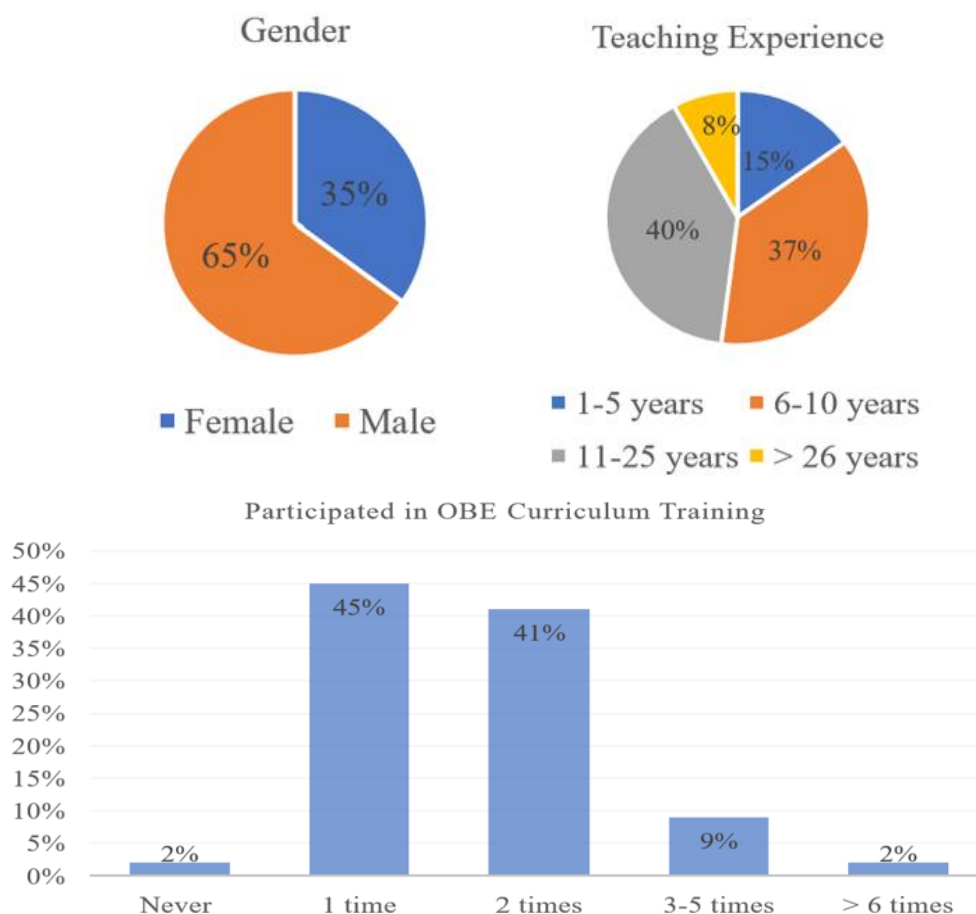
The research conducted one-sample t-tests to compare the mean attitude scores against a benchmark value of 4, which represents agreement with OBE-related statements. The benchmark value of 4 was selected because it symbolises "agree" on the five-point Likert scale. The differences among the five attitudinal dimensions were assessed using paired-sample t-tests. This test was used because the five attitudinal dimensions were assessed among the same respondents. In addition, one-way analysis of variance (ANOVA) was conducted to examine whether teaching experience and OBE training participation were associated with statistically significant differences in faculty attitudes towards OBE implementation. One-way ANOVA was appropriate for comparing mean attitude scores across multiple teaching-experience and OBE-training groups. Statistical significance was assessed at the 5% threshold.

5. Analytical Results

5.1 Respondents' Demographic Information

Figure 1

Demographic Data of Respondents



Source: Survey Data

The demographic information of all survey participants is presented in Figure 1. The proportion of male instructors was nearly twice that of female instructors, at 64.8% and 35.2%, respectively. Concerning teaching experience, 40.1% of instructors had 11-25 years of teaching experience, 36.7% had 6-10 years, 15.7% had 1-5 years, and 7.9% had more than 26 years. Regarding the number of times instructors had attended OBE training, 45% had attended OBE training once. A further 41% had attended OBE training twice, and 9% had attended OBE training three to five times. Very few participants (2%) had attended more than six times, while 2% had not participated in any OBE training.

5.2 Reliability and Validity Testing

5.2.1 Reliability Testing

The summary statistics of faculty members' attitudes toward the OBE approach are displayed in Table 1. The analysis comprises five indicators and a total of 26 items, including the mean, standard error, variance, range, number of observations, minimum value, and maximum

value. These indicators are subsequently employed in the analysis. Both the estimated standard error and variance fall within acceptable statistical bounds. Within the dataset of 267 observations, the scale ranges from 1.6 to 5. The Cronbach's alpha coefficient is frequently used to ensure that the Likert-scale indicators used in the analysis have sufficient internal consistency. The Cronbach's alpha coefficient was 0.887. This exceeds 0.70, which is considered the established threshold. This average value indicates that all variables used for the data analysis demonstrate a high-level of internal consistency.

Table 1

Summary Statistics and Reliability Testing

Indicators	Min	Max	Mean	SD	Cronbach Alpha	Items
Knowledge	2.40	5.00	3.75	0.55	0.876	5
Belief	2.00	5.00	4.00	0.42	0.853	5
Feelings	1.60	5.00	3.96	0.55	0.862	5
Readiness	2.33	5.00	4.08	0.50	0.853	6
Acceptance	3.00	5.00	4.13	0.53	0.869	5

Source: Results were obtained from statistical analyses conducted using SPSS

5.2.2 Validity Testing

The assessment of the appropriateness of the KMO values and Bartlett's test of sphericity was conducted for each part of the questionnaire. Table 2 shows that the Kaiser-Meyer-Olkin (KMO) measure of sampling Adequacy value for the "*faculty members' attitudes towards the OBE implementation competency scale*" is 0.849, indicating that the data are suitable for factor-analysis. Bartlett's test of sphericity was statistically significant ($p < 0.001$), indicating that the data in this study are also suitable for factor analysis.

Table 2

KMO and Bartlett's Test Results

KMO Value	0.849
Bartlett's Test of Sphericity	Approx. Chi-Square
	df
	P value
	725.557
	10
	< 0.001

Source: Results were obtained from statistical analyses conducted using SPSS

5.3 Faculty Attitudes Towards OBE Implementation in Engineering Education

Descriptive frequency statistics showed that 77% of instructors strongly agreed and agreed with the OBE approach. Acceptance of OBE was supported by faculty responses at a rate of 77%. Knowledge received the lowest level of agreement, while preparedness and acceptance had the highest proportions of favourable responses (Table 3). Compared with 76.7 for beliefs, 75.96 for feelings, 82.23 for readiness, and 85.30 for acceptance, the overall agreement percentage for knowledge was 64.32%. This pattern suggests that faculty members are willing to accept OBE; however, their knowledge of the concept has not developed to the same level. Lower levels of agreement were also observed for equal educational opportunities (Be10) and the time efficiency of OBE (Fe12) at the item level, further raising questions about the approach's efficacy.

Table 3*Results of Frequency of Descriptive Statistic on Attitude Towards OBE Approach*

Knowledge (Kn)	SA	A	MA	D	SD
Kn ₁	34 (12.7%)	137 (51.3%)	91 (34.1%)	5 (1.9%)	-
Kn ₂	27 (10.1%)	148 (55.4%)	89 (33.33%)	3 (1.1%)	-
Kn ₃	41 (15.4%)	141 (52.8%)	85 (31.8)	-	-
Kn ₄	34 (12.7%)	140 (52.4%)	91 (34.1%)	1 (0.4%)	1 (0.4%)
Kn ₅	30 (11.2%)	127 (47.6%)	101 (37.8%)	9 (3.4%)	-
Total =	12.42%	51.9%	34.23%	1.36%	0.08%
Beliefs (Be)					
Be ₆	68 (25.5%)	148 (55.4%)	49 (18.4%)	2 (0.7%)	-
Be ₇	81 (30.3%)	134 (50.2%)	50 (18.7%)	2 (0.7%)	-
Be ₈	79 (29.6%)	135 (50.6%)	52 (19.5%)	1 (0.4%)	-
Be ₉	40 (15%)	154 (57.7%)	69 (25.8%)	4 (1.5%)	-
Be ₁₀	61 (22.8%)	124 (46.4%)	77 (28.8%)	5 (1.9%)	-
Total =	24.64%	52.06%	22.24%	1.04%	-
Feelings (Fe)					
Fe ₁₁	52 (19.5%)	153 (57.3%)	60 (22.5%)	1 (0.4%)	1 (0.4%)
Fe ₁₂	44 (16.5%)	133 (49.8%)	85 (31.8%)	4 (1.5%)	1 (0.4%)
Fe ₁₃	63 (23.6%)	132 (49.4%)	64 (24%)	7 (2.6%)	-
Fe ₁₄	72 (27%)	152 (56.9%)	41 (15.4%)	2 (0.7%)	-
Fe ₁₅	51 (19.1%)	162 (60.7%)	51 (19.1%)	3 (1.1%)	-
Total =	21.14%	54.82%	22.56%	1.26%	0.16%
Readiness (Re)					
Re ₁₆	50 (18.7%)	155 (58.1%)	60 (22.5%)	2 (0.7%)	-
Re ₁₇	60 (25.8%)	138 (51.7%)	58 (21.7%)	1 (0.4%)	1 (0.4%)
Re ₁₈	53 (19.9%)	165 (61.8%)	47 (17.6%)	2 (0.7%)	-
Re ₁₉	51 (19.1%)	158 (59.2%)	56 (21%)	2 (0.7%)	-
Re ₂₀	81 (30.3%)	159 (59.6%)	26 (9.7%)	1 (0.4%)	-
Re ₂₁	121 (45.3%)	119 (44.6%)	27 (10.1%)	-	-
Total =	26.52%	55.83%	17.1%	0.37%	0.07%
Acceptance (Ac)					
Ac ₂₂	55 (20.6%)	161 (60.3%)	51 (19.1%)	-	-
Ac ₂₃	68 (25.5%)	163 (61%)	36 (13.5%)	-	-
Ac ₂₄	63 (23.6%)	154 (57.7%)	48 (18.4%)	-	1 (0.4%)
Ac ₂₅	72 (27%)	158 (59.2%)	37 (13.9%)	-	-
Ac ₂₆	119 (46.4%)	121 (45.3%)	27 (10.1%)	-	-
Total =	28.62%	56.7%	15%	-	(0.08%)

Source: Results were obtained from statistical analyses conducted using SPSS

5.4 The Attitude Components that Faculty Agree the Most and the Least Towards OBE Approach

There is some variation among the five attitudinal dimensions presented in Table 4. Knowledge was the lowest ($Kn = 3.75$), while acceptance was the highest ($Ac = 4.13$), followed closely by preparedness ($Re = 4.08$). Acceptance and preparedness were significantly above the benchmark value of 4, whereas knowledge was significantly below it. Beliefs, sentiment, and attitude were largely unchanged from the baseline.

Knowledge scores were significantly lower than beliefs, sentiment, preparedness, and acceptance scores, as confirmed by the paired-sample comparisons in Table 5. Conversely,

there was no significant difference between acceptance and preparedness scores. These data suggest that the primary attitudinal gap is not between willingness and acceptance but between positive attitudes towards OBE and faculty knowledge of the approach.

Table 4

Results of One Sample T-test (Test value =4)

Variables	Means	t	df	p	Conclusion
Cluster 1: Knowledge (Kn)	3.75	-7.39	266	<0.001	Different
Cluster 2: Beliefs (Be)	4.00	0.09	266	0.928	Not different
Cluster 3: Feelings (Fe)	3.96	-1.29	266	0.196	Not different
Cluster 4: Readiness (Re)	4.08	2.67	266	0.008	Different
Cluster 5: Acceptance (Ac)	4.13	4.04	266	<0.001	Different
FM's attitude toward OBE	3.98	-0.55	266	0.581	Not different

Source: Results were obtained from statistical analyses conducted using SPSS

Table 5

Results of Paired Sample T-test

Paired Samples	Means	t	df	p	Conclusion
Pair 1	(Kn=3.75) – (Be=4.00)	-9.13	266	<0.001	Different
Pair 2	(Kn=3.75) – (Fe= 3.96)	-6.42	266	<0.001	Different
Pair 3	(Kn=3.75) – (Re= 4.08)	-10.71	266	<0.001	Different
Pair 4	(Be=4.00) – (Fe= 3.96)	1.71	266	0.87	Not different
Pair 5	(Re= 4.08) – (Be=4.00)	2.92	266	0.004	Different
Pair 6	(Re= 4.08) – (Fe= 3.96)	4.86	266	<0.001	Different
Pair 7	(Ac=4.13)– (Re= 4.08)	1.99	266	0.47	Not different
Pair 8	(Ac=4.13) – (Fe=3.96)	5.65	266	<0.001	Different
Pair 9	(Ac=4.13) – (Kn=3.75)	11.79	266	<0.001	Different

Source: Results were obtained from statistical analyses conducted using SPSS

5.5 The Teaching Experience and OBE Training Affect Faculty Members' Attitude of OBE Implementation

The ANOVA results in Table 6 revealed significant differences among teaching-experience groups across all five attitude measures. The largest difference was observed for knowledge, suggesting that faculty members' understanding of OBE might be particularly influenced by their teaching experience. Differences among OBE training groups were also significant across all five variables. As before, there was considerable variation in knowledge, but differences in acceptance were quite small. These patterns suggest that OBE training and teaching experience are related to overall attitudes towards OBE, and specifically, to faculty members' understanding of OBE and perceived ability to implement it.

Table 6

The Impact of Participants' Demographic Variables on Their Attitudes

Dependent variables	Results of the One-way ANOVA	
	Teaching experiences	Obtained OBE training
Cluster 1: Knowledge (Kn)	$f(266) = 12.62, p < 0.001$	$f(266) = 12.30, p < 0.001$
Cluster 2: Belief (Be)	$f(266) = 6.20, p < 0.001$	$f(266) = 5.88, p < 0.001$

Cluster 3: Feelings (Fe)	$f(266) = 3.16, p = 0.025$	$f(266) = 9.04, p < 0.001$
Cluster 4: Readiness (Re)	$f(266) = 5.73, p < 0.001$	$f(266) = 5.06, p < 0.001$
Cluster 5: Acceptance (Ac)	$f(266) = 3.33, p = 0.020$	$f(266) = 2.58, p = 0.038$

Source: Results were obtained from statistical analyses conducted using SPSS

6. Discussion

The results reveal a gap between faculty support for OBE and their conceptual understanding of the approach. While beliefs, feelings, readiness, and acceptance tended towards strong agreement, knowledge was consistently the lowest among the dimensions surveyed. This finding supports Katawazai's (2021) conclusion that agreement can exist alongside poor understanding, but also aligns with the findings of Alimyar (2020) and Baguio (2019) who report favourable faculty opinions of OBE. Thus, it could be said that within Laos, policymakers' acceptance of OBE has outpaced faculty understanding of its principles and requirements. Nevertheless, the unusually high levels of acceptance and willingness support this perspective. Faculty members are interested in transitioning to an Outcomes-Based Education model despite their lower knowledge scores, which may indicate challenges in translating this interest into course development, proper alignment and day-to-day instruction. However, because faculty members showed less agreement that OBE saves teaching time, they might struggle with the approach's implementation. Consistent with Jansen's (1998) findings on conceptual issues, implementation may therefore be the biggest hurdle, rather than faculty disagreement with OBE. Differences between groups with varying teaching experience and groups with OBE training provide further insight into how faculty expertise influences perceptions of OBE. Significant differences in knowledge also suggest that targeted training and practice could lead to a better understanding of OBE. This aligns with Alimyar's (2020) and Mogashoa's (2013) emphasis on training and implementation capacity. Occasional workshops on OBE may not be enough for Laos' public universities; continuous training could help convert positive perceptions into effective teaching strategies and assessments.

7. Conclusion and Implications

7.1 Conclusion

The findings of the survey revealed that faculty members generally have favourable attitudes towards the implementation of outcomes-based education (OBE) in engineering programmes in Lao public universities. Regarding the beliefs, readiness, and acceptance dimensions, large percentages of faculty members reported strong agreement or agreement with statements related to OBE. This indicates institutional openness to educational change. In addition, replacing traditional content-based educational strategies with outcomes-based instruction is accepted at the attitudinal level.

However, there is a clear disparity across attitudinal dimensions. Although preparedness and acceptance levels are high, knowledge of the OBE approach remains limited. This suggests that faculty members accept and are willing to implement OBE, but do not fully understand its underlying principles and processes. Positive attitudes alone are insufficient for effective pedagogical change, and limited knowledge represents a key constraint. Overall, the results show that OBE implementation in the Lao context is characterised by strong attitudinal support, but limited conceptual and practical integration.

The quantitative analysis shows that faculty members' attitudes differ significantly based on teaching experience and participation in OBE training workshops. This indicates that

professional experience influences both knowledge and readiness to implement educational reforms. Faculty members who had more teaching experience and had participated in OBE training demonstrate stronger and more consistent positive attitudes across all dimensions. This finding suggests that professional development is central to the implementation process. Limited participation in training may explain the gap between acceptance and knowledge. These findings suggest that institutional change occurs unevenly across different dimensions of faculty readiness. While OBE is widely accepted, applying its practices to classroom teaching is constrained by limited practical experience. This highlights the need to align technical capacity with positive attitudes in curriculum design, teaching, and assessment.

These results improve our understanding of how faculty-level factors influence OBE implementation in engineering higher education in Laos. The findings indicate that acceptance, knowledge, experience, and training interact to support successful OBE implementation.

This study extends the previous research by using different statistical methods to better understand how faculty members respond to educational reform. Even if faculty members accept the concept of OBE, there can still be insufficient knowledge and capacity to implement it. The role of knowledge and training in the acceptance-implementation gap helps explain why institutional change may not correlate with classroom practice.

7.2 Policy Implications

Participation in OBE training and workshops should be one of the top priorities for policymakers and higher education administrators when planning faculty professional development. The results of this study indicate that only a small proportion of faculty members have sufficient knowledge of OBE. Therefore, professional development should be more structured, systematic, and continuous. Faculty members should be encouraged to participate in workshops that focus on OBE principles and support the application of these principles in curriculum design, teaching, and assessment.

Moreover, since teaching experience was found to significantly influence instructors' attitudes towards OBE, capacity building should be integrated into both pre-service and in-service training. This would allow faculty members to be better prepared to implement OBE in practice.

Finally, faculty members should receive institutional support to sustain OBE implementation after training. Resources such as teaching materials, feedback mechanisms, and ongoing support should be made available to faculty members. If institutions do not provide adequate support to apply the outcomes of training, positive changes in teaching practices will not be sustained. Policymakers should strengthen institutional structures to support the implementation of OBE in engineering education in Laos.

Acknowledgement

The authors would like to express their sincere gratitude and thanks to the Department of Higher Education of the Ministry of Education and Sports (MOES), the participating universities for granting permission and facilitating the study. Special thanks are extended to the Academic Affairs Departments of Faculty of Engineering of National University of Laos, Souphanouvong University, Champasack University, as well as the Academic Affairs Department of Faculty of Engineering and Academic Affairs Department of Faculty of Information and Technology of Savannakhet University for their coordination and assistance during data collection.

References

- Akhmadeeva, L., Hindy, M., & Sparrey, C. J. (2013). Overcoming obstacles to implementing an outcome-based education model: Traditional versus transformational OBE. *Proceedings of the Canadian Engineering Education Association (CEEA)*. <https://doi.org/10.24908/pceea.v0i0.4913>
- Akramy, S. A. (2021). Implementation of outcome-based education (OBE) in Afghan universities: Lecturers' voices. *International Journal of Quality in Education*, 5(2), 27–47. <https://ijqejournal.com/index.php/ijqe/article/view/27>
- Ali, Q. I. (2023). Implementing a direct assessment strategy for outcome-based education in computer engineering: A case study. *International STEM Journal*, 4(2), 24–37. <https://doi.org/10.22452/stem.vol4no2.3>
- Alimyar, Z. (2020). Outcome-based education training workshops: A study to explore their effectiveness on Afghan EFL instructors' teaching methods. *International Journal of Education and Culture*, 9(1–2), 18–30.
- Allo, M. D. G. A., Sudarsi, E. T., & Taula'bi', N. (2024). Implementation of outcome based education (OBE) principles in the curriculum of the english education study program at a higher education in Toraja. *Jurnal Onoma: Pendidikan, Bahasa dan Sastra*, 10(2), 1379–1391. <https://doi.org/10.30605/onoma.v10i2.3505>
- An, I. L. (2014). Impact of outcome-based education instruction to accountancy students in an Asian university. *Asia Pacific Journal of Education, Arts and Sciences*, 1(5), 48–52.
- ASEAN University Network. (2022). *AUN-QA assessed programmes*. <https://www.aunsec.org/discover-aun/thematic-networks/aun-qa/aun-qa-assessed-programmes>
- Baguio, J. B. (2019). Outcomes-based education: Teachers' attitude and implementation. *University of Bohol Multidisciplinary Research Journal*, 7(1), 110–127. <https://doi.org/10.15631/ubmrj.v7i1.113>
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
- Borsoto, L. D., Lescano, J. D., Maquimot, N. I., Santorce, M. J. N., Simbulan, A. F., & Pagcaliwagan, A. M. (2014). Status of implementation and usefulness of outcomes-based education in the engineering department of an Asian university. *Asian Journal of Management Science and Economics*, 2(4), 14–25. <https://www.multidisciplinaryjournals.com/wp-content/uploads/2014/02/STATUS-OF-IMPLEMENTATION.pdf>
- Canini, N. D., & Dalis, I. M. (2016). Extent of implementation and influence of outcomes-based education to civil engineering students in Misamis University, Philippines. *Journal of Multidisciplinary Studies*, 5(1), 102–124.
- Cheung, A. C. K., & Man Wong, P. (2012). Factors affecting the implementation of curriculum reform in Hong Kong. *International Journal of Educational Management*, 26(1), 39–54. <https://doi.org/10.1108/09513541211194374>
- Chowdhury, R., & Das, C. (2022). Disruptive factors in implementing outcome-based education curriculum at tertiary business education: A focus on institutional readiness. *Canadian Journal of Business and Information Studies*, 4(4), 72–85. <https://www.universepg.com/cjbis/disruptive-factors-in-implementing-outcome-based-education-curriculum-at-tertiary-business-education-a-focus-on-institutional-readiness>

- Dano, G. (2024). Strengthening outcome-based education: Capability approach perspective. *Edukasiana: Jurnal Inovasi Pendidikan*, 3(2), 143–155. https://www.researchgate.net/publication/379283681_Strengthening_outcome-based_education_Capability_approach_perspective
- Davis, M. H. (2003). Outcome-based education. *Journal of Veterinary Medical Education*, 30(3), 258–263. <https://doi.org/10.3138/jvme.30.3.258>
- Deepak, J., & Venishri, P. (2018). Outcome based education OBE trend review in engineering education. *International Journal of Trend in Scientific Research and Development*, 1(Special Issue-ICAEIT2017), 41–43. <https://doi.org/10.31142/ijtsrd19126>
- Duangdala, T., Hathalong, K., Mitthiyaphone, P., Donephondy, N., & Vongpanya, K. (2023). Tracer study (2015–2019) of the National University of Laos. *Lao Social Science Journal*, 1(1), 50–58. <https://doi.org/10.71026/lss.2023.010108>
- Giessen-Hood, C. (1999). *Teachers' attitudes towards the implementation of outcomes-based education (OBE) in South Africa*.
- Government of the Lao People's Democratic Republic. (2021, September 16). *Decree on the Lao Qualifications Framework* (Decree No. 566/GOV).
- Harden, R. M., Crosby, J. R., & Davis, M. H. (1999). AMEE Guide No. 14: Outcome-based education: Part 1—An introduction to outcome-based education. *Medical Teacher*, 21(1), 7–14. <https://doi.org/10.1080/01421599979969>
- Hegde, R. (2014). Implementing outcome based education for microcontrollers. In *2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE)* (pp. 245–249). <https://doi.org/10.1109/MITE.2014.7020281>
- Isa, C. M. M., Saman, H. M., Tahir, W., Jani, J., & Mukri, M. (2017). Understanding of outcome-based education (OBE) implementation by civil engineering students in Malaysia. In *2017 IEEE 9th International Conference on Engineering Education (ICEED)* (pp. 96–100). IEEE. <https://doi.org/10.1109/ICEED.2017.8251172>
- Jadhav, M. R., Kakade, A. B., Jagtap, S. R., & Patil, M. S. (2020). Impact assessment of outcome-based approach in engineering education in India. *Procedia Computer Science*, 172, 791–796. <https://doi.org/10.1016/j.procs.2020.05.113>
- Jansen, J. D. (1998). Curriculum Reform in South Africa: a critical analysis of outcomes-based education. *Cambridge Journal of Education*, 28(3), 321–331. <https://doi.org/10.1080/0305764980280305>
- Katawazai, R. (2021). Implementing outcome-based education and student-centered learning in Afghan public universities: the current practices and challenges. *Heliyon*, 7(5), e07076. <https://doi.org/10.1016/j.heliyon.2021.e07076>
- Khanna, R., & Mehrotra, D. (2019). The roadmap for quality improvement from traditional through competency based (CBE) towards outcome based education (OBE) in dentistry. *Journal of Oral Biology and Craniofacial Research*, 9(2), 139–142. <https://doi.org/10.1016/j.jobcr.2019.02.004>
- Lavi, R., & Bagiati, A. (2022). The new engineering education transformation program at Massachusetts Institute of Technology: The evolving design and implementation of a programmatic evaluation study. In B. R. Moser, P. Koomsap, & J. Stjepandić (Eds.), *Transdisciplinarity and the future of engineering* (pp. 658–667). IOS Press. <https://doi.org/10.3233/ATDE220698>
- Loreto, M. M. (2018). Outcomes based teaching and learning practices in the hotel and resort management programme of Dusit Thani College. *Dusit Thani College Journal*, 12(Special), 82–98. <https://so01.tci-thaijo.org/index.php/journaldtc/article/view/129493>

- Macatangay, A. O., Braza, L. D., Gamboa, M. N., Gonzales, A. D., Fuentes, R. A. P., Macalalad, J. A., Hernandez, K. T., Montejo, J. R. B., & Mendoza, F. M. (2016). Status of implementation and usefulness of outcomes-based education in customs administration program of one Asian university. *Asia Pacific Journal of Education, Arts and Sciences*, 3(3), 62–69.
- Magenuka, T. K., & Sibanda, M. (2023). Developing a curriculum for the fourth industrial revolution in electrical engineering education. *Proceedings of the 10th Focus Conference (TFC 2023)* (pp. 284–298). https://doi.org/10.2991/978-2-38476-134-0_19
- Ministry of Education and Sports. (2011). *National curriculum standards for associate degree, bachelor degree, master degree, doctoral degree, first degree of specialization and second degree of specialization*.
- Ministry of Education and Sports. (2020, October 2). ຂໍ້ຕົກລົງວ່າດ້ວຍມາດຕະຖານຫຼັກສູດແຫ່ງຊາດ ລະດັບປະລິນຍາເອກ (ສະບັບປັບປຸງ) [Agreement on the national curriculum standards for doctoral degree programmes (revised edition)] (Agreement No. 4485/MOES).
- Mogashoa, T. I. (2013). The experiences of adult learning centre educators in implementing outcomes-based assessment. *Mediterranean Journal of Social Sciences*, 4(14), 455–462. <https://doi.org/10.5901/mjss.2013.v4n14p455>
- Ortega, R. A. A., & Ortega-Dela Cruz, R. A. (2016). Educators' attitude towards outcomes-based educational approach in English second language learning. *American Journal of Educational Research*, 4(8), 597–601. <https://pubs.sciepub.com/education/4/8/3/index.html>
- Pradhan, D. (2021). Effectiveness of outcome-based education (OBE) toward empowering the students' performance in an engineering course. *Journal of Advances in Education and Philosophy*, 5(2), 58–65. <https://doi.org/10.36348/jaep.2021.v05i02.003>
- Puerto-López, K. C., Lozano Cárdenas, F. J., Medina Delgado, B., Niño Rondón, C. V., & Castro Casadiego, S. A. (2023). Analysis of learning outcomes in engineering programs. *Ingeniería Solidaria*, 19(3), 1–23. <https://doi.org/10.16925/2357-6014.2023.03.07>
- Sawant, P. (2016). Implementation of outcome based education: A beginning. *Journal of Engineering Education Transformations*, 29(Special Issue), Article 85643. <https://doi.org/10.16920/jeet/2016/v0i0/85643>
- Spady, W. G. (1994). *Outcome-based education: Critical issues and answers*. American Association of School Administrators.
- Sunra, L., Aeni, N., & Sally, F. H. S. (2024). A comprehensive exploration of outcome-based education principles and practices. *Asian Journal of Education and Social Studies*, 50(1), 1–9. <https://doi.org/10.9734/ajess/2024/v50i11234>
- Yasmin, M., & Yasmeen, A. (2021). Viability of outcome-based education in teaching English as a second language to chemical engineering learners. *Education for Chemical Engineers*, 36, 100–106. <https://doi.org/10.1016/j.ece.2021.04.005>
- Yusof, R., Othman, N., Norwani, N. M., Ahmad, N. L. B., & Jalil, N. B. A. (2017). Implementation of outcome-based education (OBE) in accounting programme in higher education. *International Journal of Academic Research in Business and Social Sciences*, 7(6), 1186–1200. <https://doi.org/10.6007/IJARBS/v7-i6/3352>
- Yusoff, Y. M., Fuaad, N. F. A., Yasin, R. B. M., & Tawil, N. M. (2014). Achievement of the programme outcomes in outcomes-based education implementation: A meta-analysis. In *Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management* (pp. 1739–1749). IEOM Society. <https://ieomsociety.org/ieom2014/pdfs/388.pdf>

NATURE OF ALGEBRA TASKS IN PHILIPPINE TEXTBOOKS FOR GRADES 9 AND 10: CONTENT TOPICS

Michael G Domanais^{1*}

Ryan V Dio^{1,2}

¹*Bicol University Graduate School, Legazpi City, Philippines*

²*Sorsogon State University, Sorsogon City, Philippines*

<michael.domanais@bicol-u.edu.ph>

ABSTRACT

The mathematical tasks in the textbooks influence students' appreciation of concepts, including how they acquire the set learning competencies as directed by teachers' intentions. This descriptive study examined the nature of 301 mathematical tasks in Philippine textbooks for Grades 9 and 10, focusing on the Algebra (AL) topics in terms of content, context, and mathematical processes. The study employed content analysis of the Philippine Mathematics textbooks for Grades 9 (Department of Education, 2014) and Grade 10 (Department of Education, 2015), following the four coding stages: Coding Protocol Development, Initial Coding, Code Verification, and Coding Analysis, involving the researchers (Initial coder) and five mathematics expert coders (Coding Validators). The coding analysis reveals a limited real-life context among the review and learning tasks of the Patterns and Algebra (AL) topics, which can create opportunities for classroom innovation and differentiation by combining action-based (applied), image-based (visual), and language-based (verbal) tasks that can stimulate mathematics thinking, processes, and positive response among learners. Classroom mathematics teachers should be creative enough to choose activities that will provide students with ample exposure to all types of problems.

Keywords: Mathematics Teacher Education Research, Mathematical Tasks, Algebra, Real and Non-Real-Life Context, Mathematics Processes

1. Introduction

Mathematical tasks are fundamental to students' engagement with mathematical ideas, as they shape opportunities for developing conceptual understanding, reasoning, and problem-solving skills (Özgeldi & Esen, 2010). Embedded within instructional materials, these tasks influence how curriculum intentions are translated into classroom practice and how learners construct mathematical knowledge (Fatimah & Prabawanto, 2020; Mastuti et al., 2022). Within the Philippine K to 12 Mathematics Curriculum, AL is a particularly important content area because it develops algebraic reasoning, representation, and generalisation skills that underpin learning across other domains of mathematics (Department of Education 2016; Dio, 2020). Despite the adoption of the spiral progression approach and curriculum reforms, the consistently low mathematics performance of Filipino learners in international assessments such as PISA suggests the need to examine the learning opportunities provided through instructional resources, particularly the mathematical tasks presented in textbooks (Summaries, 2018; Golla & Reyes, 2020).

This need is especially relevant because government-issued textbooks remain the primary instructional resource in Philippine public schools and are intended for use for at least five years (Department of Education. 2012). The Grade 9 and Grade 10 mathematics textbooks, first published in 2014 and 2015, respectively, continue to be widely used despite the absence

of substantial revisions. Given the central role of textbooks in mediating curriculum implementation and shaping students' learning experiences (Valverde et al., 2002), it is timely to examine the mathematical tasks contained within the AL content area. Such an investigation can provide insights into the content, context, and cognitive demands of these tasks and generate evidence to inform future textbook revision, curriculum enhancement, and instructional improvement efforts (Johnson et al., 2020; Mastuti et al., 2023).

2. Literature Review

Textbooks play a central role in mathematics education by bridging the gap between the intended curriculum and the curriculum enacted in classrooms (Thompson, 2013). They provide a comprehensive representation of mathematical content and tasks, shaping both teaching and learning processes (Valverde et al., 2002; Houang & Schmidt, 2008). As the primary instructional resource for teachers and students, textbooks influence lesson planning, classroom activities, homework, and instructional decision-making (Bakken & Andersson-Bakken., 2021). Given their substantial influence on learning opportunities, the quality and nature of mathematical tasks embedded within textbooks are critical to achieving curriculum goals and improving mathematics achievement. This importance is further highlighted by the global learning crisis reported by the UNESCO Institute for Statistics (UNESCO Institute for Statistics [UIS], 2017), which found that more than 617 million children and adolescents fail to attain minimum proficiency levels in reading and mathematics. Similar concerns have been reflected in international large-scale assessments such as PISA and TIMSS, where the nature of mathematical tasks has been identified as a significant factor influencing student achievement (Johnson et al., 2020; Mastuti & Prayitno, 2023).

In the Philippines, concerns regarding mathematics learning are compounded by challenges in curriculum implementation, limited instructional resources, and gaps between the K to 12 Mathematics Curriculum and international standards such as the PISA Mathematics Literacy Framework (Golla & Reyes, 2020). Previous studies have identified deficiencies in contextualisation, higher-order reasoning, critical thinking, and problem-solving within the current curriculum. Although the K to 12 Mathematics Curriculum encompasses five content domains such as AL, Numbers and Number Sense, Measurement, Geometry, and Probability and Statistics (Republic Act No. 10533, 2013), AL warrants particular attention because it serves as the foundation for advanced mathematical thinking and supports reasoning, generalisation, modelling, and problem-solving across other mathematical domains. Furthermore, algebraic competence is strongly associated with future success in higher-level mathematics and STEM-related disciplines. However, evidence suggests that secondary students frequently struggle with algebraic concepts, often focusing on procedural manipulation rather than conceptual understanding and real-world application, resulting in persistent misconceptions, errors, and learning difficulties (Al-Rababaha et al., 2022). Given the pivotal role of AL in developing mathematical proficiency, it is essential to examine how this domain is represented in school textbooks. Filipino learners are expected to develop representational fluency, simplify algebraic expressions and equations, and apply appropriate strategies in solving problems (DepEd, 2016; Dio, 2020). Such outcomes require mathematical tasks that promote active engagement, conceptual understanding, and the development of 21st-century competencies, including critical thinking, adaptability, collaboration, and innovation (Chu et al., 2016). Yet previous evaluations of Philippine mathematics textbooks have reported inconsistencies with curriculum expectations, limited attention to affective and behavioural objectives, and inadequate integration of local contexts and values education (de Guzman & Adamos, 2020). Consequently, a focused analysis of the mathematical tasks related to AL in currently used secondary mathematics textbooks is necessary to determine whether they provide meaningful learning opportunities that support

curriculum objectives, address learners' difficulties, and foster the competencies expected in both national and international educational frameworks.

3. Methodology

This study employs a descriptive research design using content analysis to provide an objective account of mathematical tasks. Content analysis is a systematic and objective research method for analysing textual or visual data (Krippendorff, 2019) that adheres to the scientific principles of objectivity, systematicity, and generality (Neuendorf, 2017). The study comprehensively describes various aspects, including the nature of the content, context, and mathematical processes of a specific set of mathematical tasks in the textbooks, focusing on AL content topics. To ensure objectivity, the research utilised defined rules that allow for consistent interpretation and replication of findings across different observers. Systematicity is maintained by applying standardised criteria for categorising content, ensuring a rigorous and consistent selection process. Furthermore, the study achieved generality by connecting these observations to established theoretical frameworks, moving beyond simple description to provide broader scientific value.

3.1 Participant-Coders

Seven participants were involved in this study, divided into primary and secondary coders. The researchers served as the primary coders, responsible for determining the codes for each mathematical task based on the study's analytical framework. The remaining five participants served as secondary coders, tasked with verifying the initial codes assigned by the researchers. These secondary coders were selected based on specific criteria: they must have conducted research in mathematics education, and either currently serve as master teachers who attended the National Training of Trainers or Regional Mass Training for Mathematics 9 and 10, or work as college instructors/professors specialising in mathematics courses.

3.2 Sources of Data

This study employed content analysis on the widely circulated DepEd public school such as the 1st edition of the mathematics textbooks for Grade 9 and Grade 10, both written in English and structured into four units composed of modules and lessons. Each lesson follows a standardised format containing four distinct task sections: "*What to Know*" (recalling prior knowledge on the various mathematics concepts and skills previously learned and link them to the lesson to be studied), "*What to Process*" (applying new principles), "*What to Reflect and Understand*" (reflecting on and summarising concepts), and "*What to Transfer*" (practicing and assessing real-world application).

Table 1 summarises the total number of mathematical tasks in the Grades 9 and 10 textbooks review, learning, and practice/assessment of each AL content topic that was subjected to content analysis. A total of 301 AL tasks, representing 225 or 74.75% of the total tasks in the Grade 9 textbook, and 76 or 25.25% of the total were included in the Grade 10 textbook.

Table 1*Number of AL Tasks Analysed*

Mathematical Tasks	Textbook		Over-all	
	Grade 9	Grade 10	Total	%
Review Task	60	10	70	23.26
Learning Task	85	50	135	44.85
Practice/Assessment Task	80	16	96	31.89
Total	225	76	301	100

3.3 Data Gathering Procedures

Guided by a literature review following *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) guidelines, a coding manual for mathematics content, context, and processes was developed to systematically analyse the nature of mathematical tasks in Grades 9 and 10 textbooks across four procedural stages: Coding Protocol Development, Initial Coding, Code Verification, and Coding Analysis.

The content nature of mathematical tasks defines the type of information they present, coded as Purely Mathematical (PM) for tasks containing numerical, algebraic, or symbolic representations, Verbal (Vb) for text-based descriptions, Visual (Vs) for graphical, tabular, or diagrammatic representations, and Combination (Cb) for tasks integrating multiple formats, such as mathematical-verbal (M-Vb), mathematical-visual (M-Vs), verbal-visual (Vb-Vs), or all three (M-Vb-Vs) (Hidayah & Forgasz, 2020; Yang & Sianturi, 2022).

The context nature of mathematical tasks characterises their real-world applicability (Almuna Salgado, 2016; Summaries, 2018), categorising tasks as either non-real context for abstract algorithmic exercises or real-life context across four codes: personal (individual/family/peer activities), occupational (work/professional), societal (community/governance/national issues), and scientific (natural world/technology/intra-mathematical).

The process nature of mathematical tasks defines the cognitive and problem-solving actions students utilise during task completion (Erbilgin, 2017), aligned with Common Core Mathematics Standards across eight codes: Making Sense of Problems and Persevering in Solving Them (PS), Reasoning Abstractly and Quantitatively (RAQ), Constructing Viable Arguments and Critiquing Reasoning (CACRO), Modelling with Mathematics (M), Using Tools Strategically (UT), Attending to Precision (AP), Looking for and Using Structure (MUS), and Expressing Regularity in Repeated Reasoning (ER).

The initial coding process was conducted by the primary coders over three weeks—allocating one week each to the content, context, and mathematical process dimensions—with adjustments made through collaborative brainstorming. To verify these classifications, two independent secondary coders per dimension were onboarded through individual orientations utilising a manual equipped with concrete, illustrative examples. Any subsequent coding discrepancies between the primary researcher and secondary coders were systematically resolved using a strict consensus rule, accepting only the codes that achieved at least a 2-out-of-3 rater match ($\geq 67\%$ agreement).

Finally, the coding analysis stage established data reliability and minimised individual inaccuracy through rigorous inter-rater corroboration. Inter-rater reliability was evaluated using percentage agreement—the most popular consensus estimate for assessing classification consistency across nominal coding units (Stemler, 2004). Percent agreement is calculated as *Percentage Agreement = (Number of agreed-upon codes / Total number of codes assigned) × 100%*. In content analysis, consensus estimates operate under the principle that when independent judges come to an exact agreement on qualitative categories, they demonstrate a shared conceptual understanding of the underlying construct (Stemler, 2004).

3.4 Data Analysis Procedures

Guided by Miles and Huberman (1994), the data analysis seamlessly encompassed data reduction, data presentation, and conclusion drawing and verification. Qualitative task details were condensed and transformed into percentage-based tables and graphs, allowing the researcher to iteratively cross-examine data displays, raw data, and existing literature to produce accurate, evidence-backed findings.

4. Results and Discussion

4.1 Nature of AL Tasks in Terms of Content

The nature of the content of the AL tasks in the Philippine government-issued textbooks for Grades 9 and 10, along with review tasks, learning tasks, and practice/assessment tasks, can either be PM, Vb, or Cb as shown in Table 2. The majority of tasks, whether review, learning, or practice/assessment, fall into the Cb form.

Table 2

Distribution of Mathematical Tasks in Grades 9 and 10 Textbooks by Nature of Content

Grade Level	Mathematical Task	Nature of Content						
		Purely Math (PM)	Visual (Vs)	Verbal (Vb)	Combination (Cb)			
					M-Vb	M-Vs	Vb-Vs	M-Vb-Vs
9	Review	3.33%	0.00%	5.00%	45.00%	1.67%	35.00%	10.00%
	Learning	5.88%	0.00%	10.59%	37.65%	1.18%	12.94%	31.76%
	Practice/Assessment	0.00%	0.00%	32.50%	20.00%	0.00%	28.75%	18.75%
10	Review	40.00%	0.00%	0.00%	20.00%	0.00%	10.00%	30.00%
	Learning	32.00%	0.00%	0.00%	44.00%	0.00%	2.00%	22.00%
	Practice/Assessment	6.25%	0.00%	25.00%	43.75%	0.00%	12.50%	12.50%

Table 2 further reveals that within the Grade 9 Mathematics textbook, review tasks are primarily characterised by the integration of multiple modes, specifically the PM-Vb and Vb-Vs combinations at 45.00% and 35.00%, respectively. Learning tasks maintain a similar pattern, with PM-Vb accounting for 37.65% and PM-Vb-Vs for 31.76%. In contrast, practice/assessment tasks exhibit a shift toward linguistic and spatial representations, Vb tasks constituting 32.50%, while 28.75% are Vb-Vs.

Regarding the Grade 10 Mathematics textbook, a large proportion at 40.00% of the review tasks are PM in nature. This emphasis on symbolic representation persists through the Learning tasks, where M-Vb (44.00%) and PM (32.00%) comprise the majority of the content. However, consistent with the patterns identified in Grade 9, Grade 10 practice/assessment tasks transition away from pure abstraction toward integrated representations, with 43.75% utilising PM-Vb combinations and 25.00% relying on Vb nature.

The PM tasks in these textbooks were strategically designed for procedural fluency and algorithmic practice. Typically presented as drills, these activities require learners to recall what they have previously studied from previous levels or from previous topics within the current grade level and help them improve their accuracy and speed in solving algorithmic exercises. According to Adams and Maki (2023), drill-based interventions have been demonstrated to be effective in enhancing learners' foundational academic skills. Thus, providing these kinds of drills to elicit prior knowledge will help the learners recall and master the prerequisite skills.

Conversely, the Vb tasks in these textbooks were designed to focus on everyday language. These tasks often appear as word problems about real-life situations. Furthermore, other tasks require learners to come up with their own real-world examples. This method is particularly significant as Hwang et al. (2025) suggest that while problem-posing is a powerful student-centred approach for improving both cognitive and affective learning, it is frequently underrepresented in traditional textbooks. By incorporating these prompts, the textbooks help learners think more deeply about how mathematics connects to the world around them. This not only makes the learning process more relevant and interesting but also builds student agency and ownership, as it allows them to express their own mathematical inquiries and interests.

The mathematical tasks that are a Cb can either be M-Vb, M-Vs, Vb-Vs, or M-Vb-Vs. These multi-modal tasks were designed to bridge the gap between concrete illustrations and abstract algebraic concepts. M-Vb tasks often consist of a numerical drill followed by “guide questions. This combination ensures that learners do not just perform calculations blindly but also reflect on and justify their mathematical process. An example of a common M-Vs task in the textbooks requires learners to write an equation based on a provided geometric figure. This helps learners see the physical meaning behind the algebraic variables. Meanwhile, Vb-Vs tasks use pictures and texts together, such as a diagram of a real-world object accompanied by a story or a set of instructions. This supports iconic and enactive learning by providing a visual scaffold for the language used in the problem.

Nirawati et al. (2020) suggest that to overcome the abstract nature of symbols, Algebra requires learners to have multi-representation ability, as using more than one representation during problem-solving is linked to higher student scores. The result highlights that these textbooks can offer a spectrum of learning opportunities, encompassing action-based (enactive), image-based (iconic), and language-based (symbolic) tasks. Therefore, it can be inferred that while these textbooks lack pure visual tasks, they are still pedagogically valuable because its focus on combination tasks allows students to develop the essential skill of linking verbal, visual, and symbolic information to analyse conditions and solve problems.

Figure 1

Sample mathematical task for Grade 9 that is combination in nature

Activity 18. Create a Problem for Me?
Formulate a problem based on the illustration and its corresponding attribute. Show your solution and final answer for the created problem. Your work shall be evaluated according to the rubric

Given: the time taken by light to travel 1 meter is roughly 3×10^{-8} seconds

Formulated Problem and Solution:



Source: Mathematics 9 Textbook (Department of Education, 2014)

4.2 Nature of AL Tasks in Terms of Context

Table 3 illustrates the distribution of mathematical tasks in both textbooks, categorised by the nature of their contexts. The contexts were categorised into distinct categories – non-real, personal, occupational, societal and scientific contexts. The coding results indicate that one mathematical task may involve one or more context types. The results indicate that while all context types are represented in varying proportions, the majority of mathematical tasks regardless of whether they are for review, initial learning, or assessment utilise non-real-world contexts.

Table 3

Distribution of Mathematical Tasks in Grades 9 and 10 Textbooks by Nature of Context

Grade Level	Mathematical Task	Type of context				
		Non-Real	Real Context			
			Personal	Occupational	Societal	Scientific
9	Review	39	5	8	1	10
	Learning	60	8	8	1	19
	Practice/ Assessment	29	13	28	11	29
10	Review	8	1	0	0	1
	Learning	32	1	0	0	17
	Practice/ Assessment	5	3	5	2	6

Reflecting a broader global pattern of prioritising abstract mathematical instruction (Wijaya et al., 2015; Yang et al., 2017), a majority of tasks in both textbooks (Table 4)—including identifying quadratic equations and filling in formal definitions—operate strictly in non-real settings centred on skill practice and symbolic manipulation. Grounded in cognitive theory,

this non-contextual focus assumes that learners at the formal operations stage can process abstract mathematical concepts without concrete, real-world framing (Domingo, 2023). Nevertheless, such exercises remain vital for strengthening abstract reasoning, procedural fluency, and conceptual understanding necessary for real-world applications (National Council of Teachers of Mathematics, 2000).

Table 4

Distribution of Mathematical Tasks in Grades 9 and 10 Textbooks by Nature of Context

Grade Level	Mathematical Task	Type of context				
		Non-Real	Real Context			
			Personal	Occupational	Societal	Scientific
9	Review	39	5	8	1	10
	Learning	60	8	8	1	19
	Practice/ Assessment	29	13	28	11	29
10	Review	8	1	0	0	1
	Learning	32	1	0	0	17
	Practice/ Assessment	5	3	5	2	6

Source: *Mathematics 9 Learner's Material* (Department of Education, 2014) and Source: *Mathematics 10 Learner's Module* (Department of Education, 2015)

Alongside these abstract exercises, both textbooks actively incorporate diverse real-world contexts, including personal, occupational, societal, and scientific domains, to bridge abstract concepts with authentic applications. Personal contexts engage learners through familiar daily experiences such as sports, physical recreation, arts and crafts, and personal finance management. Occupational contexts simulate real-world professional constraints; for instance, the Grade 9 “*Design Packaging Boxes*” activity requires students to formulate quadratic equations to design packaging under specific volume parameters, modelling work in industrial and packaging engineering. Societal contexts address community issues, local landmarks, recycling campaigns, and health awareness initiatives like HIV awareness aligned with Sustainable Development Goals (SDGs) to demonstrate how mathematics tracks social progress. Finally, scientific contexts connect mathematical ideas to physics, chemistry, and astronomy. Examples include Grade 9 tasks where students formulate word problems around atomic diameters, electron charges, and light travel times, enabling them to interpret computations within physical frameworks. As Yang et al. (2017) assert, embedding tasks in these meaningful real-life scenarios facilitates a vital shift from rote symbolic manipulation to authentic application, ensuring learners comprehend both the procedural execution and the underlying rationale of their mathematical work.

4.4 Nature of AL Tasks in Terms of Mathematical Processes

Figure 2 summarises the distribution of the mathematical processes involved in the AL review, learning, and practice/assessment tasks for Grades 9 and 10 textbooks. At the final stage of code verification, there were a total of 446 codes assigned to different mathematical tasks in the mathematics 9 textbook and 110 codes in the mathematics 10 textbook.

It can be gleaned from Figure 2 that mathematical tasks in both textbooks provide opportunities for learners to engage with eight identified mathematical processes though the emphasis on each process varies across different mathematical tasks. However, it can be noted that the occurrence varies across different types of mathematics tasks.

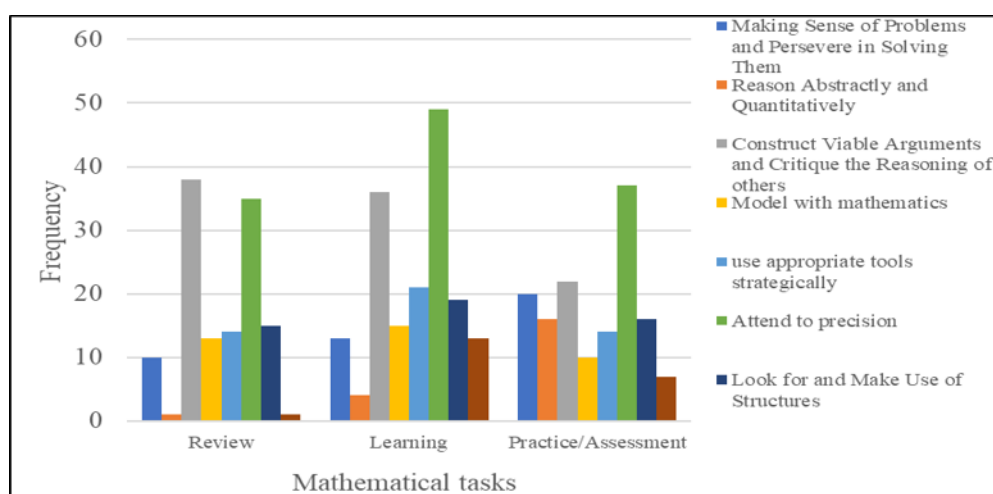
4.4.1 Making Sense of Problems and Persevering in Solving Them (PS)

Figure 2 illustrates a progressive increase in the frequency of tasks targeting PS across review, learning, and practice/assessment phases. The observed progression of PS transitioning from context-grounded review tasks to autonomous creation and assessment reflects an intentional curricular design aimed at fostering sustained mathematical engagement. By embedding and systematically increasing the integration of PS activities across all three instructional phases, the learning sequence provides structured problem-solving exposure that normalises perseverance as a fundamental learning behaviour rather than treating sense-making as an isolated event (Daniel & Dinapoli, 2024). In Grade 9, this process adheres to a structured instructional continuum. Review tasks ground new concepts in familiar, real-world contexts such as “*Activity 6: Let’s Shoot that Ball!*” (Module 1, Lesson 3), where students compute time and distance in basketball motion while evaluating the physical feasibility of the ball reaching 160 ft, thereby bridging prior knowledge with emerging algebraic ideas.

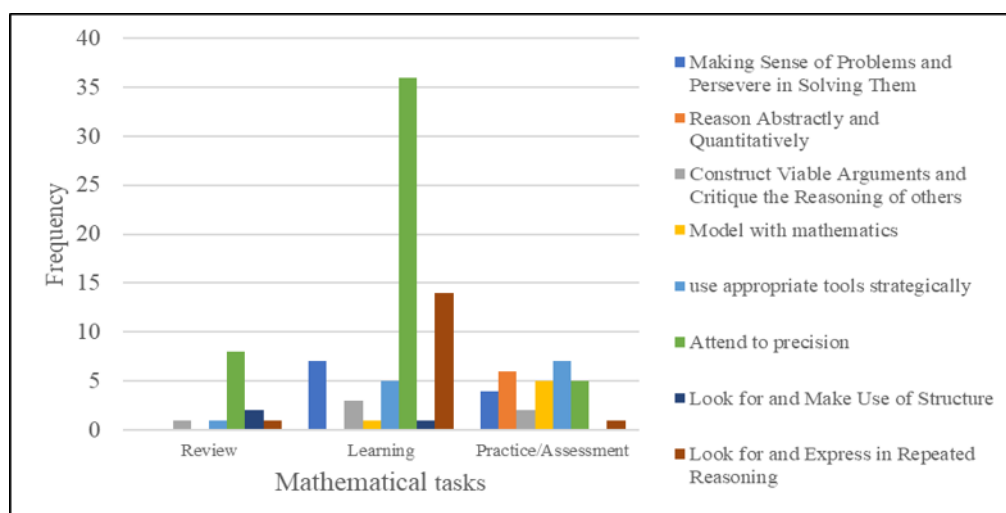
A closer analysis of this sequence reveals an intentional shift from initial conceptual retrieval to high-level mathematical autonomy. As learners transition into main learning tasks, PS focuses on model formulation and procedural fluency; for instance, *Activity 8: Let’s Make a Table!* (Module 1, Lesson 3) challenges students to translate verbal statements into quadratic equations and justify dimensional validity without direct computation, cultivating a deeper conceptual understanding of quadratic roots. This developmental arc culminates in the practice and assessment tasks, where PS requires synthesis, justification, and mathematical creation. As demonstrated in “*Activity 10: Will It or Will It Not?*” (Module 1, Lesson 3), learners progress beyond solving pre-formulated models to constructing their own contextual problems, deriving the corresponding quadratic expressions, and solving them. This shows a progression that transitions students from guided sense-making to independent problem-solving mastery.

Figure 2

Nature of Mathematical Processes in Grades 9 and 10: AL Tasks



Note. Grade 9 Mathematics Textbook



Note. Grade 10 Mathematics Textbook

4.4.2 Unit Reasoning Abstractly and Quantitatively (RAQ)

Figure 2 demonstrates that specific mathematical tasks were designed to develop RAQ as a core mathematical process. According to Erbilgin (2017), RAQ tasks demand learners to reason abstractly by representing a given situation symbolically and make sense of the formula, and to reason quantitatively by making sense of quantities in real-world contexts, estimating quantities, comparing quantities using units, examining relationships between quantities, and making inferences from data. In Grade 9, this process was introduced through tasks demanding symbolic representation and conceptual sense-making such as analysing the zero-product property in “*Activity 3: Why is the product zero?*” (Module 1, Lesson 2B) and the discriminant in “*Activity 10: Investigate me!*” (Module 1, Lesson 7) as well as tasks assessing physical feasibility, such as determining if a mathematical root logically corresponds to a physical side length in “*Activity 7: What does a square have?*” (Module 1, Lesson 2A).

By Grade 10, the RAQ process progresses to more complex, multi-step projects like the “*Reality Series*” (Activity 14, Module 1, Lesson 1), where learners identify empirical trends (e.g., financial growth), decontextualise the data into tabular and algebraic models, and subsequently re-contextualise their mathematical findings through peer explanations of real-world mechanics. By embedding these tasks, the textbooks move beyond rote procedural execution to position mathematics as an analytical tool for real-world problem solving, thereby cultivating the critical thinking competencies required for mathematical proficiency (Kilpatrick et al., 2001; Sachdeva & Eggen, 2021).

4.4.3 Constructing Viable Arguments and Critiquing the Reasoning of Others (CACRO)

In Grade 9, the CACRO process serves as the secondary focal point in the textbook, emphasising the development of high-order discursive skills. The prevalence of CACRO is evidenced by the “Guide Questions” integrated into nearly all tasks. For instance, prompts such as “*Compare your answer with those of your classmates. Did you get the same answer? If not, explain,*” and “*Compare your work with those of other groups,*” explicitly evoke the critiquing of others' reasoning. This suggests that while computational precision is the primary objective, the textbook places a significant weight on communicative competence. By engaging in tasks designed to cultivate CACRO, learners are required to articulate the underlying logic of mathematical procedures, analyse misconceptions, and evaluate diverse strategies. Such requirements move beyond rote memorisation, necessitating that students

actively participate in substantive mathematical discourse and demonstrate a meta-cognitive understanding of their work and the work of others. These findings support Ray's (2020) assertion that mathematical tasks must provide a diversity of critique opportunities, as a lack of variety in assessment can constrain the breadth of students' mathematical reasoning.

4.4.4 Modelling with Mathematics (M)

The mathematical process, model with mathematics (M), is embedded in the mathematical tasks in the textbooks. Rather than merely solving given equations, learners are required to represent situations symbolically by translating these scenarios into mathematical language (Haines et al., 2000). For example, in Grade 9, "*Activity 3: Air Out!*" (Module 1, Lesson 2A) asks students to draw a diagram, define variables, and formulate an equation to solve a construction problem, while "*Activity 10: Will It or Will It Not?*" (Module 1, Lesson 3) uses a model of a basketball's path to examine relationships between quantities like time and height.

In Grade 10, modelling as a process is seen through tasks on recognising patterns and solving complex geometric problems by making sense of formulas. For instance, in "*Activity 4: What Do We Have in Common?*" (Module 1, Lesson 1), learners use matchsticks to identify growth patterns and come up with a mathematical model for the sequence, fostering their ability to create and interpret mathematical models that describe growth patterns. Similarly, in "*Activity 14: Modelling through Polynomial Equations*" (Module 2, Lesson 3), learners are challenged to set up a polynomial equation model and solve it. By grounding these tasks in real-world scenarios like installing an exhaust fan or shooting a basketball, the textbooks provide what Berget (2022) describes as the essential starting points for the modelling process. Each task requires students to mathematicise a situation by translating verbal narratives into expressions and making inferences from data. However, Berget (2022) also warns that if students only encounter highly structured and simplified tasks, they may miss out on the full experience of applying mathematics to the messy, unstructured reality of life. Erbaş et al. (2014) assert that integrating modelling into mathematics education is important for improving students' problem-solving and analytical thinking abilities. Thus, the need for more teaching and learning resources integrating this mathematical process.

4.4.5 Using Appropriate Tools Strategically (UT)

Figure 2 shows that UT is among mathematical processes with high frequency. This means that the mathematical tasks provide learners with opportunities to engage in selecting and using appropriate materials for solving mathematical problems or investigating mathematical ideas. In Mathematics 9, several tools were used - either single tool or multiple tools to enhance understanding and problem-solving skills. These include making use of tables to show relationships between figures and blocks, cartesian planes to graph the relations and visual observations through graphics/figures. In Mathematics 10, the tools used were similar to those in Mathematics 9 textbook. Research shows that providing strategic tools improves mathematical problem-solving skills, though effectiveness depends heavily on individual student background and context (Cooper & Alibali, 2012).

4.4.6 Attending to Precision (AP)

Figure 2 illustrates that across both textbooks, the highest proportion of tasks, whether for review, learning, or practice/assessment, is designed for the AP process. This indicates that the tasks were strategically designed to ensure learners develop the discipline to utilise formal mathematical terminology, labels, and notations; describe situations with appropriate mathematical language, perform processes fluently, and calculate efficiently and correctly. Collectively, this focus on AP reflects an instructional goal of bridging conceptual understanding with technical exactness, ensuring that learners can convey mathematical

ideas with clarity while maintaining computational reliability. Engledowl et al. (2015) highlights that AP is a vital mathematical process since it prevents errors in logic and helps learners transition from basic arithmetic to advanced algebraic reasoning. It further allows learners to express and refine mathematical ideas, construct meaning, and become active participants in mathematical discourses rather than passive recipients of knowledge.

4.4.7 Looking for and making use of mathematical structures (MUS)

The textbooks also provide opportunities for learners to deal with MUS. Küchemann (2010) highlights the value of helping students advance from basic pattern recognition to a more powerful, generic understanding of mathematical structure.

In Mathematics, various mathematical tasks involve analysing numbers to classify them as real, rational, or perfect squares and analysing equations to determine if they are quadratic inequalities. One of the examples of a review task, “*Activity 1: Which are Real? Which are Not?*” (Module 1, Lesson 3), was designed to help learners classify and understand the properties of various numbers, such as rational and irrational values. Another learning task, “*Activity 4: Quadratic Inequalities or Not?*” (Module 1, Lesson 7) focuses on analysing mathematical sentences to identify and classify quadratic inequalities and, practice/assessment task, “*Activity 16: Mathematics in Parabolic Bridges!*” (Module 2 Lesson 3) uses the GRASP strategy to demonstrate the real-world application of math by modelling the curves of actual bridges using quadratic equations.

Mathematics 10, on the other hand, involves analysing expressions to identify non-polynomials and analysing polynomial functions. “*Activity 1: Spot the Difference*” (Module 2, Lesson 1) tasks students to identify expressions that are not polynomials by examining their structure. They classify the expressions as polynomials or not based on their characteristics (e.g. presence of negative exponents, radicals or division by variables). This task helps them to connect their understanding of polynomials to broader algebraic concepts. “*Activity 2: Fix and Move Them, then Fill Me Up*” (Module 3, Lesson 1) prompts the learners to analyse the given polynomial functions and their structures. They classify parts of the given polynomial functions according to their properties – degree, leading coefficient, constant term.

4.4.8 Expressing Regularity in Repeated Reasoning (ER)

The textbooks also provide opportunities for learners to deal with ER. It is present in mathematical tasks, though with varying frequencies for each grade level and is most prevalent in learning tasks compared to review and practice/assessment tasks. For example, in “*Activity 1: Describe me in many ways*” (Module 2, Lesson 1), learners engage in a multi-layered exploration of patterns, moving from the visual observation and drawing of succeeding figures to the formalisation of mathematical relationships. This progression requires students to translate patterns into tables, ordered pairs, and Cartesian graphs while identifying key variables such as domain, range, dependent and independent variables. By synthesising these various representations to derive an appropriate equation, students move beyond simple computation to a comprehensive mastery of repeated reasoning (Mulligan & Mitchelmore, 2009).

5. Conclusion

The analysis of Grade 9 and Grade 10 mathematics textbooks highlights how specific choices in task design shape opportunities for learning. The books regularly combine words, pictures, and symbols within tasks. This mix helps students see ideas in different ways, though the main goal of many tasks remains applying a step-by-step formula. Similarly, while tasks include

real-world settings like everyday personal situations or workplace examples, most problems supply all the required data directly within the prompt. This gives students practice applying methods in familiar setups, though it leaves less room for open-ended problem-solving such as deciding which information matters or building a model from scratch, among other skills. Practicing these open-ended skills is particularly helpful for international assessments like PISA and TIMSS, which frequently use unfamiliar scenarios. Because textbook tasks serve different learning goals, teachers play an important role in adjusting them, such as by adding support when a problem is too difficult or asking extra questions to stretch students when a problem is too straightforward. These findings offer curriculum developers practical ideas for balancing routine practice with more open-ended questions in future editions.

These findings also suggest focused directions for future research. Classroom-based observational studies can examine how teachers enact specific textbook tasks in real time, assessing the extent to which task demand changes during instruction. Additionally, comparative analyses benchmarking local textbook tasks against specific sample units from high-performing jurisdictions on international assessments can identify structural differences in task complexity. Finally, empirical research into student cognitive strategies when navigating transitions between symbolic, visual, and narrative representations can further inform evidence-based textbook design.

Acknowledgment

Originating from the first author's PhD dissertation at Bicol University under the second author's supervision, this study was funded and supported by the Department of Science and Technology–Science Education Institute (DOST-SEI) through its Capacity Building Program for Science and Mathematics Education.

References

- Adams, S. R., & Maki, K. E. (2023). Differential effectiveness, efficiency, and acceptability of drill-based interventions targeting multiplication facts. *Psychology in the Schools*, 60, 2520–2539. <https://doi.org/10.1002/pits.22880>
- Almuna Salgado, F. (2016). *Investigating the impact of context on students' performance*. Mathematics Education Research Group of Australasia. <https://eric.ed.gov/?id=ED572407>
- Al-Rababaha, Y., Yew, W. T., & Meng, C. C. (2022). Algebra misconceptions among tenth graders in the United Arab Emirates. *International Journal of Academic Research in Progressive Education and Development*, 11(1), 713–728. <https://doi.org/10.6007/ijarped/v11-i1/12197>
- Bakken, J., & Andersson-Bakken, E. (2021). The textbook task as a genre. *Journal of Curriculum Studies*, 53(6), 729–748. <https://doi.org/10.1080/00220272.2021.1929499>
- Berget, I. K. L. (2022). Mathematical modelling in textbook tasks and national examination in Norwegian upper secondary school. *NOMAD Nordic Studies in Mathematics Education*, 27(1). <https://www.semanticscholar.org/paper/Mathematical-modelling-in-textbook-tasks-and-in-Berget/78931d3b7c8219fde5414d3b363bbbd9a5f93cdb>
- Chu, S. K. W., Reynolds, R. B., Tavares, N. J., Notari, M., & Lee, C. W. Y. (2016). *21st century skills development through inquiry-based learning: From theory to practice*. Springer.

- Cooper, J. L., & Alibali, M. W. (2012). Visual representations in mathematics problem-solving: Effects of diagrams and illustrations. North American Chapter of the International Group for the Psychology of Mathematics Education. <https://files.eric.ed.gov/fulltext/ED584910.pdf>
- Daniel, A., & DiNapoli, J. (2024). Investigating pedagogies in undergraduate precalculus and their relationships to students' attitudes towards mathematics and perseverance in problem-solving. *Journal of Research in Science, Mathematics and Technology Education*, 7(SI), 61–83. <https://doi.org/10.31756/jrsmte.314SI>
- de Guzman, A. B., & Adamos, J. L. (2020). Like the layers of an onion: Curricular noticing as a lens to understand the epistemological features of the Philippine K to 12 secondary mathematics curriculum materials. *Educational Research for Policy and Practice*, 19, 389–409. <https://doi.org/10.1007/s10671-020-09264-8>
- Department of Education. (2012). *Policy and guidelines on the proper distribution, care, recording, retrieval and disposal of textbooks (TXs) with the teacher's manuals (TMs) and other instructional materials* (DepEd Order No. 14, s. 2012). <https://www.deped.gov.ph/2012/02/10/do-14-s-2012-policy-and-guidelines-on-the-proper-distribution-care-recording-retrieval-and-disposal-of-textbooks-txs-with-the-teachers-manuals-tms-and-other-instructional-materials/>
- Department of Education. (2014). *Mathematics 9 Learner's Material* (First Edition).
- Department of Education. (2015). *Mathematics 10 Learner's Module* (First Edition).
- Department of Education. (2016). K to 12 Mathematics Curriculum Guide
- Dio, R. V. (2020). Exploring vertical coherence of content topics in Philippine spiral Kto10 mathematics curriculum. *International Journal of Learning, Teaching and Educational Research*, 19(11), 259-282. <https://doi.org/10.26803/ijlter.19.11.15>
- Domingo, J. G. (2023). Determinants of cognitive skills achievement in mathematics of elementary pre-Service teachers. *Journal for Educators, Teachers and Trainers*, 14(5), 254-273. <https://doi.org/10.47750/jett.2023.14.05.025>
- Engledowl, C., Otten, S., & Spain, V. (2015). The discourse of attending to precision in secondary classrooms. North American Chapter of the International Group for the Psychology of Mathematics Education. <https://files.eric.ed.gov/fulltext/ED584240.pdf>
- Erbaş, A. K., Kertil, M., Çetinkaya, B., Çakıroğlu, E., Alacacı, C., & Baş, S. (2014). Mathematical modeling in mathematics education: Basic concepts and approaches. *Educational Sciences: Theory & Practice*, 14(4), 1621–1627. <https://doi.org/10.12738/estp.2014.4.2039>
- Erbilgin, E. (2017). A comparison of the mathematical processes embedded in the content standards of Turkey and Singapore. *Research in Social Sciences and Technology*, 2(1). <https://doi.org/10.46303/ressat.02.01.3>
- Fatimah, A. T., & Prabawanto, S. (2020). Mathematical understanding and reasoning of vocational school students in agriculture-based mathematical tasks. *Journal for the Education of Gifted Young Scientists*, 8(2), 771-782. <https://doi.org/10.17478/jegys.702884>
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academies Press. <https://doi.org/10.17226/9822>
- Golla, E. F., & Reyes, A. G. (2020). PISA mathematics literacy framework vis-à-vis the Philippine K to 12 mathematics curriculum. In *Challenges of PISA: The PNU report* (p. 57). Philippine Normal University.
- Haines, C., Crouch, R., & Davis, J. (2000). *Mathematical modeling skills: A research instrument* (Technical Report No. 55). University of Hertfordshire, Faculty of Engineering and Information Sciences. <https://doi.org/10.13140/RG.2.2.22034.45763>

- Hidayah, M. & Forgasz, H. (2020). A comparison of mathematical tasks types used in Indonesian and Australian textbooks based on geometry contents. *Journal on Mathematics Education*, 11(3), 385-404. <https://eric.ed.gov/?id=EJ1294881>
- Houang, R. T., & Schmidt, W. H. (2008, September 18–20). *TIMSS international curriculum analysis and measuring educational opportunities* [Conference presentation]. 3rd IEA International Research Conference. <https://www.iea.nl/publications/presentations/timss-international-curriculum-analysis-and-measuring-educational>
- Hwang, S., Han, J., Muirhead, F., et al. (2025). Textbooks as a resource for teaching mathematics through problem posing: Catalyzing instructional change. *ZDM—Mathematics Education*, 57, 987–1001. <https://doi.org/10.1007/s11858-025-01694-z>
- Johnson, H. L., McClintock, E. D., & Gardner, A. (2020). Opportunities for reasoning: Digital task design to promote students' conceptions of graphs as representing relationships between quantities. *Digital Experiences in Mathematics Education*, 6(3), 340–366. <https://doi.org/10.1007/s40751-020-00061-9>
- Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications. <https://doi.org/10.4135/9781071878781>
- Küchemann, D. (2010). Using patterns generically to see structure. *Pedagogies: An International Journal*, 5(3), 233–250. <https://doi.org/10.1080/1554480X.2010.486147>
- Mastuti, A. G., & Prayitno, L. L. (2023). Exploring high school teacher's design of rich algebra tasks. *Jurnal Elemen*, 9(1), 1-14. <https://doi.org/10.29408/jel.v9i1.5851>
- Mastuti, A. G., Abdillah, & Rijal, M. (2022). Teachers promoting mathematical reasoning in tasks. *JTAM (Jurnal Teori dan Aplikasi Matematika)*, 6(2), 371–385. <https://doi.org/10.31764/jtam.v6i2.7339>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage Publications, Inc.
- Mulligan, J., & Mitchelmore, M. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal*, 21(2), 33–49. <https://doi.org/10.1007/BF03217544>
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.
- Neuendorf, K. A. (2017). *The Content Analysis Guidebook* (2nd ed.). <https://doi.org/10.4135/9781071802878>
- Nirawati, R., Darhim, S. Fatimah, & A. Irma. (2020). Analysis of multi-representation ability to solve algebra problem. *Journal of Physics: Conference Series*, 1657, Article 012075. <https://doi.org/10.1088/1742-6596/1657/1/012075>
- Özgeldi, M., & Esen, Y. (2010). Analysis of mathematical tasks in Turkish elementary school mathematics textbooks. *Procedia - Social and Behavioral Sciences*, 2(2), 2277–2281. <https://doi.org/10.1016/j.sbspro.2010.03.322>
- Ray, A. (2020). What does it mean to critique? An exploration of the types of critiques promoted in seventh-grade mathematics curriculum and assessment materials. *School Science and Mathematics*, 120(5), 273–284. <https://doi.org/10.1111/ssm.12406>
- Republic Act No. 10533. (2013, May 15). *Enhanced basic education act of 2013*. Official Gazette of the Republic of the Philippines. <https://officialgazette.gov.ph/2013/05/15/republic-act-no-10533/>

- Sachdeva, S., & Eggen, P.-O. (2021). Learners' critical thinking about learning mathematics. *International Electronic Journal of Mathematics Education*, 16(3), Article em0644. <https://doi.org/10.29333/iejme/11003>
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation*, 9(1), Article 4. <https://doi.org/10.7275/96jp-xz07>
- Summaries, C. E. (2018). PISA 2018 Results.
- Thompson, D. R. (2013). Reasoning-and-proving in the written curriculum: Lessons and implications for teachers, curriculum designers, and researchers. *International Journal of Educational Research*, 64, 141–148. <https://doi.org/10.1016/j.ijer.2013.09.013>
- UNESCO Institute for Statistics. (2017). *More than one-half of children and adolescents are not learning worldwide* (Fact Sheet No. 46). UNESCO Institute for Statistics.
- Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Springer. <https://doi.org/10.1007/978-94-007-0844-0>
- Wijaya, A., van den Heuvel-Panhuizen, M., & Doorman, M. (2015). Opportunity-to-learn context-based tasks provided by mathematics textbooks. *Educational Studies in Mathematics*, 89, 41–65. <https://doi.org/10.1007/s10649-015-9595-1>
- Yang, D. C., & Sianturi, I. A. J. (2022). Analysis of algebraic problems intended for elementary graders in Finland, Indonesia, Malaysia, Singapore, and Taiwan. *Educational Studies*, 48(1), 75–97. <https://doi.org/10.1080/03055698.2020.1740977>
- Yang, D.-C., Tseng, Y.-K., & Wang, T.-L. (2017). A comparison of geometry problems in middle-grade mathematics textbooks from Taiwan, Singapore, Finland, and the United States. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(7), 2841–2857. <https://doi.org/10.12973/eurasia.2017.00721a>

DEVELOPING AND VALIDATING A MINECRAFT-BASED COMPUTATIONAL THINKING TEACHING MODEL FOR CODING EDUCATION IN PRIMARY SCHOOLS

Goh Kok Ming

SJKC Hua Lian 1, Malaysia

<kokming888@gmail.com>

ABSTRACT

Computational thinking is increasingly recognised as a core capability in primary coding education, yet novice learners often struggle to connect visible programming actions with abstract concepts such as variables, conditionals, operators, testing and debugging. Minecraft Education offers a visual, construction-based coding environment, but prior research suggests that game-based engagement alone is insufficient for deep computational understanding. This study developed and validated the Minecraft Computational Thinking Teaching Model (MCT-TM), a pedagogical model designed to guide learners from concrete Minecraft construction towards abstract computational reasoning. A design-based research approach was used across four phases: literature-based needs analysis, model design, expert validation using the Fuzzy Delphi Method (FDM), and model refinement. Seven experts in educational technology, coding education, computational thinking, curriculum design, assessment, and digital learning reviewed the model. All ten components met the FDM acceptance criteria, with threshold values from 0.068 to 0.151, expert consensus from 85.71% to 100.00%, and fuzzy scores from 0.83 to 0.93. The validated model consists of eight cyclical stages: Explore, Use, Predict, Investigate, Modify, Debug, Create, and Reflect. Its novelty lies in integrating UMC, PRIMM, and ABR into a Minecraft-specific structure that explicitly connects pedagogical action, computational thinking concepts, and assessment evidence. The study focuses on theoretical and expert validity, where classroom effectiveness requires future empirical implementation.

Keywords: Minecraft Education, Computational Thinking, Coding Education, Game-Based Learning

1. Introduction

Computational thinking (CT) has become a significant learning outcome in primary education because students are expected not only to use digital technologies but also to understand how digital solutions are designed, tested and improved (Van Schoors et al., 2024). Coding education provides a practical route for CT development because learners must sequence instructions, design algorithms, test solutions, debug errors and create computational artefacts (Kutay & Öner, 2022). However, novice learners often experience coding as abstract and cognitively demanding, especially when concepts such as coordinates, variables, conditionals and operators are introduced without sufficient concrete support.

Minecraft Education is frequently used to support coding because it provides an open-ended, visual and construction-based environment. Students can immediately see how code changes the game world: a loop may produce repeated block placement, an event may trigger Agent movement, and a coordinate error may place a structure in an unexpected location. These affordances make Minecraft promising for young learners who need concrete experience before formal abstraction (Bile, 2022; Mørch & Andersen, 2023). Nevertheless, the

educational value of Minecraft depends on pedagogy. Students may enjoy building and exploring without being able to explain the computational logic behind their artefacts.

Existing coding pedagogies offer useful but incomplete support for Minecraft coding. Use-Modify-Create (UMC) scaffolds progression from guided use to independent creation. PRIMM encourages learners to predict, run, investigate, modify, and make programmes. Action-Breakdown-Repair (ABR) frames debugging as a productive learning process. Yet these models were not originally designed to address Minecraft's distinctive spatial, visual and artefact-based coding affordances. A Minecraft-specific model is therefore needed to connect gameplay, code comprehension, modification, debugging, artifact creation, reflection, and assessment into a coherent instructional sequence.

This study responds to this need by developing and validating the Minecraft Computational Thinking Teaching Model (MCT-TM). The study addresses two research questions: (i) What pedagogical principles should inform a Minecraft-based CT teaching model for primary coding education? (ii) To what extent do experts agree on the relevance, clarity, practicality, pedagogical coherence, CT alignment, and classroom usability of the proposed model? The contribution of the study is conceptual and methodological: it offers an expert-validated model for future classroom implementation, but it does not claim evidence of student impact at this validation stage.

2. Literature Review

2.1 Coding Education and Computational Thinking

Computational Thinking (CT) in coding education involves more than producing correct code. It includes decomposition, algorithmic reasoning, abstraction, testing, debugging, iteration, and explanation of design decisions. Brennan and Resnick's (2012) framework is useful because it distinguishes computational concepts, computational practices and computational perspectives. This distinction is important for Minecraft coding: a student may build a visually impressive artefact but still require support to explain the concepts, practices, and reasoning behind it. Empirical evidence shows both the promise and the limitation of Minecraft-based coding. Kutay and Öner (2022) reported improvements in learners' CT knowledge through Minecraft coding activities, but also found uneven application of concepts. More visible concepts such as sequences, loops, events, and parallelism were applied more successfully than less visible concepts such as variables, operators, and conditionals. This suggests that Minecraft should not be treated as self-sufficient pedagogy. Learners require structured opportunities to predict code behaviour, investigate code-output relationships, modify programmes, debug errors and reflect on computational meaning.

2.2 Minecraft Education as a Game-Based Coding Environment

Minecraft Education aligns with game-based learning because it provides goals, interaction, immediate feedback, and meaningful play. It also aligns with constructionism because learners create shareable artefacts that can be tested, discussed, and improved. These features can support motivation and creativity, but they can also distract from formal learning if building is not connected to explicit computational reasoning. Slattery et al. (2025) concluded that Minecraft has broad educational potential, yet the evidence base remains methodologically limited. Therefore, model development should be carefully conducted, as Minecraft may enable CT learning, but claims of effectiveness require rigorous classroom studies. The key theoretical issue is the transition from concrete construction to abstract reasoning. Mørch and Andersen (2023) showed that Minecraft programming can move learners from block building and modelling towards higher-level functions and general-

purpose programming. However, students may struggle when abstraction increases. The proposed MCT-TM addresses this transition by making Predict, Investigate, Debug, and Reflect explicit stages rather than optional teacher activities.

2.3 Game-Based Learning and Constructionism in Minecraft Coding

Minecraft-based coding is often associated with game-based learning because it provides opportunities for goal-oriented problem-solving, experimentation, collaboration, and immediate feedback. These features can support engagement and learning, and studies have reported positive outcomes in areas such as creativity, critical thinking, and academic achievement. However, the evidence remains inconclusive because much of the existing research is affected by methodological limitations and risks of bias (Slattery et al., 2025). As a result, Minecraft's educational value should not be assumed solely based on its popularity or engaging nature. From a constructionist perspective, Minecraft enables students to create meaningful digital artefacts through code, allowing abstract computational concepts to be expressed in tangible forms (Nkadimeng & Ankiewicz, 2022). Nevertheless, the creation of impressive artefacts does not necessarily indicate a deep understanding of the underlying computational logic. Students may achieve successful outcomes through trial-and-error, imitation, or peer assistance without developing authentic computational thinking skills (Kutay & Öner, 2022; Giannakoulas & Xinogalos, 2024). This appearance–reality gap suggests that Minecraft-based coding activities require structured pedagogical support, including reflection, debugging, and explanation of coding decisions, to ensure that meaningful learning accompanies artifact creation.

2.4 Concrete-to-Abstract Progression in Minecraft Coding

Minecraft is often promoted as an effective tool for helping students progress from concrete experiences to abstract computational thinking because it allows learners to begin with visible building activities before engaging with increasingly complex coding tasks. Research suggests that Minecraft can make abstract mathematical, geometrical, and programming concepts more accessible by connecting them to tangible actions within a familiar environment (Bile, 2022). Similarly, Mørch and Andersen (2023) proposed a progression from manual construction to automated building through functions and eventually to general-purpose programming concepts. While this progression provides a useful framework for scaffolding learning, it assumes that engagement with coding activities will naturally lead to deeper conceptual understanding, an assumption that requires critical examination. Past studies suggest that movement across abstraction levels is not automatic. Although students may successfully build structures and use coding tools such as MakeCode, Python, or JavaScript, they often struggle to explain the computational principles underpinning their solutions, particularly when dealing with variables, parameters, coordinates, and other abstract concepts (Mørch & Andersen, 2023). As a result, the production of successful artefacts may create a misleading impression of computational understanding, even when learners possess only superficial knowledge of the code involved (Kutay & Öner, 2022; Giannakoulas & Xinogalos, 2024). This limitation highlights the need for explicit instructional support that encourages students to investigate code-output relationships, debug errors, and reflect on their computational reasoning.

2.5 Existing Teaching Models Relevant to Minecraft Coding

Several pedagogical models provide useful scaffolding for Minecraft-based coding education, particularly Use–Modify–Create (UMC), PRIMM, and Action–Breakdown–Repair (ABR). UMC supports a gradual transition from interacting with existing code to modifying and ultimately creating original programmes, thereby reducing cognitive load for novice learners. PRIMM extends this progression by emphasising code comprehension before code production,

encouraging students to predict, discuss, and investigate programme behaviour before making changes (Sentance & Waite, 2017). These approaches align well with Minecraft because learners can observe the immediate effects of code on a visible game world. However, both models largely assume that engagement with code naturally leads to understanding, despite evidence that students may successfully manipulate or adapt code without fully grasping the computational concepts involved.

The ABR model addresses this limitation by highlighting the educational value of breakdowns and debugging, viewing errors as opportunities for deeper learning rather than failures (Mørch & Andersen, 2023). Nevertheless, ABR focuses primarily on problem resolution and does not provide a comprehensive instructional sequence for coding education. Consequently, each model contributes valuable but partial support for Minecraft-based learning. UMC promotes progression toward creation, PRIMM strengthens code comprehension, and ABR emphasises debugging and conceptual development. Their individual limitations suggest the need for a more integrated approach. Therefore, the proposed MCT-TM combines these complementary elements through the stages of Explore and Use, Predict and Investigate, Modify and Debug, Create, and Reflect, to support not only artifact construction but also deeper computational thinking and metacognitive understanding.

2.6 Research Gap and Need for the Proposed Model

The reviewed literature suggests that Minecraft Education offers significant potential for coding education by providing immediate visual feedback, supporting creative construction, facilitating collaboration, and enabling progression from block-based to text-based programming (Suwannik, 2022; Voštinár & Dobrota, 2022). These affordances appear well suited to the development of computational thinking, particularly when students engage in artifact creation, experimentation, and debugging (Kutay & Öner, 2022). However, much of the existing literature focuses on engagement and learning outcomes rather than on the pedagogical mechanisms by which learners transition from game-based activities to abstract computational reasoning. As a result, while Minecraft is frequently presented as an effective learning environment, there remains limited understanding of how specific instructional processes contribute to computational thinking development (Mørch & Andersen, 2023; Slattery et al., 2025).

A further limitation is that existing pedagogical models, such as UMC, PRIMM, and ABR, offer valuable but fragmented forms of scaffolding and are rarely integrated into a coherent Minecraft-specific framework (Mørch & Andersen, 2023). Moreover, despite promising findings, the overall evidence base remains methodologically constrained, limiting confidence in claims about educational effectiveness (Slattery et al., 2025). These gaps indicate a need for theoretically grounded and systematically validated instructional models. Therefore, this study proposes the Minecraft Computational Thinking Teaching Model (MCT-TM) as an integrated approach that synthesises principles from game-based learning, constructionism, concrete-to-abstract programming, UMC, PRIMM, and ABR (Bile, 2022; Mørch & Andersen, 2023). Through the stages of Explore, Use, Predict, Investigate, Modify, Debug, Create, and Reflect, the model aims to support students' progression from concrete Minecraft-based activities to deeper computational reasoning and understanding.

3. Methodology

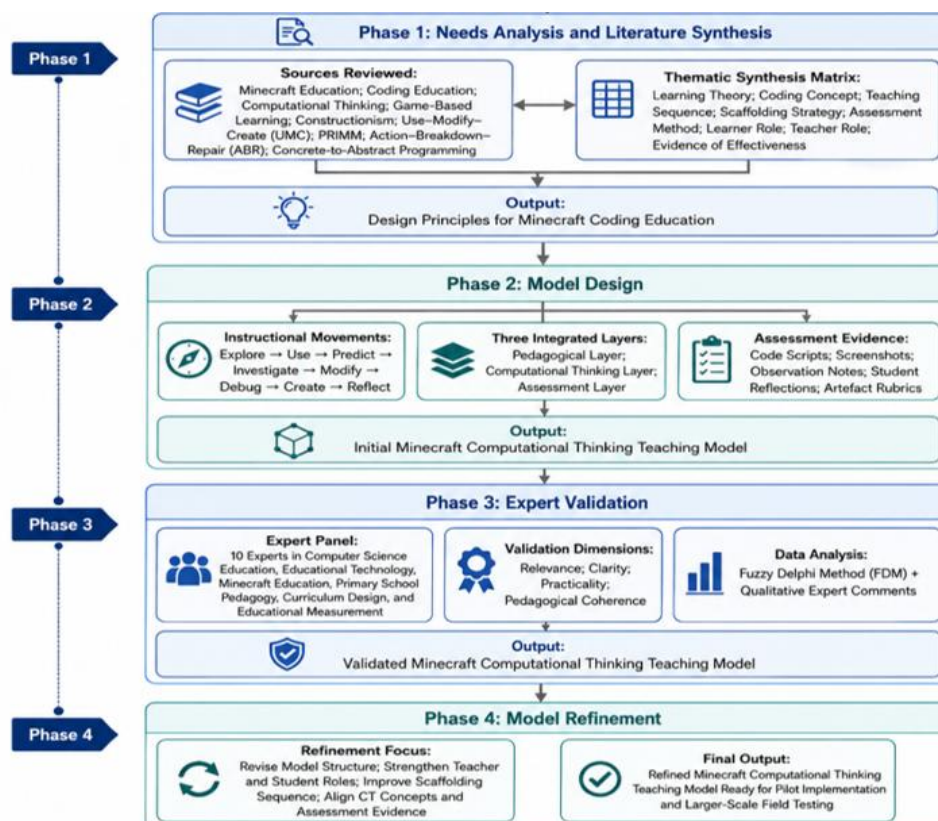
3.1 Research Design

This study used design-based research because the aim was to develop, validate, and refine a pedagogical model rather than test an existing intervention. Four phases were conducted,

as shown in Figure 1: (i) needs analysis and literature synthesis, (ii) initial model design, (iii) expert validation using FDM, and (iv) refinement based on quantitative consensus and qualitative expert feedback. Classroom implementation and student outcome evaluation are positioned as future research.

Figure 1

Research Design: The Flow of the Study



3.2 Phase 1: Needs Analysis and Literature Synthesis

Past literature on Minecraft Education, coding education, CT, game-based learning, constructionism, concrete-to-abstract programming, UMC, PRIMM, and ABR was reviewed using a synthesis matrix. Each source was examined according to learning theory, coding concepts, teaching sequence, scaffolding strategy, assessment method, learner role, teacher role, and evidence of effectiveness. The synthesis generated design principles: begin with concrete Minecraft challenges, use existing code before independent creation, connect code to visible output, treat debugging as learning, support progression towards abstraction, provide differentiated pathways, and use multiple assessment sources.

3.3 Phase 2: Model Design

The second phase involved constructing the initial version of the Minecraft Computational Thinking Teaching Model. The model was designed by integrating the design principles generated from Phase 1 with established coding pedagogy models. The initial model consisted of eight stages: Explore, Use, Predict, Investigate, Modify, Debug, Create, and Reflect. These stages were arranged as a cycle rather than a one-way lesson plan because authentic coding involves repeated testing, modification, and repair. The model was organised into three layers: a pedagogical layer describing teacher and learner activity; a CT layer mapping tasks to concepts and practices; and an assessment layer identifying evidence of learning.

3.4 Phase 3: Expert Validation

The third phase involved validating the initial MCT-TM using the Fuzzy Delphi Method (FDM). Seven experts were invited using purposive selection. Experts were required to meet at least two of the following criteria: (i) academic qualification or professional expertise in education or computing; experience in educational technology, coding education, CT, curriculum design, assessment or Minecraft Education; (ii) experience with primary or school-level digital learning; and (iii) publication, training, module development or implementation experience related to digital pedagogy. The panel included academics, teacher educators, experienced digital learning practitioners, and curriculum or assessment specialists. To protect confidentiality, individual identities are not reported. Experts evaluated the model using a structured questionnaire covering relevance, clarity, practicality, pedagogical coherence, CT alignment, and classroom usability. A seven-point linguistic scale was converted into triangular fuzzy numbers. FDM was used because it captures consensus under conditions of judgemental uncertainty. Components were accepted when three criteria were met: threshold value $d \leq 0.20$, expert consensus $\geq 75\%$, and fuzzy score indicating an acceptable level of agreement. Qualitative comments were analysed thematically to identify refinements related to differentiation, reflection, debugging, assessment, and teacher guidance.

3.5 Phase 4: Model Refinement

The fourth phase involved refining the MCT-TM based on the FDM results and expert feedback. The refinement process focused on improving the clarity, practicality, and coherence of the model. Revisions were made to the descriptions of the instructional stages, the mapping between Minecraft activities and computational thinking concepts, the role of debugging, the assessment evidence, and the visual representation of the model. The refined model retained the eight instructional stages of Explore, Use, Predict, Investigate, Modify, Debug, Create, and Reflect. However, expert feedback led to clearer explanations of how teachers could scaffold students during each stage. The refinement also strengthened the assessment layer by making explicit the use of multiple sources of evidence, including coding artefacts, observation checklists, student reflections, and pre- and post-tests. The final output of this phase was a validated and refined Minecraft Computational Thinking Teaching Model for future classroom implementation, pilot testing, and empirical evaluation of its impact on students' computational thinking development.

3.6 Ethical Considerations

Experts received information about the purpose of the study, procedures, voluntary participation, confidentiality, and responsible data handling. Responses were reported in aggregate form. Because this stage involved expert validation only, no claims are made about student learning outcomes or classroom effectiveness.

4. Findings

4.1 Design Principles and Initial Model

The first phase identified key pedagogical principles from past literature, where Minecraft can support coding education when its game-based affordances are connected to computational thinking concepts and practices. The findings showed that Minecraft coding should be planned as more than game participation. Effective use requires scaffolding from concrete challenges to code reasoning, from code use to modification and creation, and from visible artefacts to explicit CT explanation. These principles informed the eight-stage model shown in Table 1.

Table 1*Design Principles Derived from the Literature Synthesis*

Design Principle	Justification	Implication For MCT-TM
Begin with concrete Minecraft challenges	Novice learners understand abstract programming ideas more easily when they first experience them through visible and meaningful actions.	The model begins with Explore to situate learning in a Minecraft problem context.
Use existing code before asking students to create original code	Use–Modify–Create and PRIMM suggest that learners benefit from reading, running, and understanding code before independent creation.	The Use, Predict, and Investigate stages are placed before Modify and Create.
Connect the code to the visible output	Minecraft allows students to see how commands, coordinates, loops, and events produce visible outcomes in the game world.	The Investigate stage focuses on code-output analysis.
Treat debugging as a learning process	Coding involves testing, error identification, repair, and refinement. Debugging supports computational thinking practices.	Debug is included as a distinct instructional stage rather than a hidden activity.
Support progression from concrete to abstract reasoning	Students need scaffolding to move from block building to code-based construction and later to general computational concepts.	The model links Minecraft artefacts to CT concepts such as sequence, loop, event, data, condition, and abstraction.
Provide differentiated pathways	Students enter Minecraft coding with different readiness levels, from block-based beginners to text-based coders.	Beginner, intermediate, and advanced pathways are added to the refined model.
Use multiple assessment sources	Computational thinking cannot be assessed only through correct code; artefacts, reflections, observations, and interviews are also needed.	The assessment layer includes pre/post-test, artefact rubric, observation checklist, student reflection, and interview.

The findings suggested that the proposed model should not treat Minecraft as merely a motivational game environment. Instead, Minecraft should be positioned as a structured coding environment where students construct artefacts, test ideas, identify errors, explain computational concepts, and reflect on their coding decisions. Table 2 explains the initial structure of the proposed Minecraft Computational Thinking Teaching Model.

Table 2*Initial Structure of the Minecraft Computational Thinking Teaching Model*

Model Stage	Teacher Role	Student Role	CT Focus	Evidence of Learning
Explore	Introduces a Minecraft challenge	Explores problem context	Decomposition, problem orientation	Initial idea, discussion, sketch
Use	Provides sample code	Runs the given code	Sequence, event	Working code output
Predict	Asks guiding questions	Predicts code outcome	Algorithmic reasoning	Prediction notes
Investigate	Guides code-output analysis	Compares code and result	Pattern recognition, abstraction	Code explanation
Modify	Provides scaffolded changes	Changes code parameters	Loop, data, operator	Modified artefact
Debug	Guides error analysis	Identifies and fixes errors	Testing and debugging	Corrected code

Create	Facilitates project work	Designs original artefact	Abstraction, modularisation	Minecraft coding project
Reflect	Supports metacognition	Explains coding decisions	CT practices and perspectives	Reflection log or presentation

In summary, the findings suggested that the proposed model is not a linear coding lesson sequence but a cyclical learning model. Students transit repeatedly between the stages, such as Modify, Debug, and Create, as they test and improve their work.

4.2 Expert Validation Using FDM

All ten evaluated components met the FDM acceptance criteria. Threshold values were below 0.20, consensus percentages exceeded 75%, and fuzzy scores were acceptable. Debug achieved the highest fuzzy score, indicating strong expert support for making debugging a visible instructional process rather than an incidental technical activity. Reflect, Differentiated Pathway and Assessment Layer obtained slightly lower scores, but they remained accepted and were refined based on expert comments. Table 3 shows the FDM findings for the main components of the model.

Table 3

FDM Findings for Main Model Components

Model Component	Threshold Value (d)	Expert Consensus (%)	Fuzzy Score	Decision
Explore	0.083	100.00	0.91	Accepted
Use	0.071	100.00	0.92	Accepted
Predict	0.126	85.71	0.86	Accepted
Investigate	0.089	100.00	0.90	Accepted
Modify	0.076	100.00	0.92	Accepted
Debug	0.068	100.00	0.93	Accepted
Create	0.072	100.00	0.92	Accepted
Reflect	0.134	85.71	0.85	Accepted
Differentiated pathway	0.142	85.71	0.84	Accepted
Assessment layer	0.151	85.71	0.83	Accepted

Based on Table 3, the highest fuzzy score was recorded for Debug (0.93), followed by Use, Modify, and Create (0.92). This suggests that experts strongly valued the active coding, experimentation, artefact development, and error-correction elements of the model. The high ranking of Debug is especially meaningful because it shows that experts viewed debugging not as a technical failure, but as a central computational thinking practice in Minecraft-based coding education. On the other hand, the Reflect, the Differentiated pathway, and the Assessment layer received slightly lower fuzzy scores, although all remained within the accepted range. These lower scores did not indicate rejection. Instead, they suggested that these components required clearer operationalization. Experts agreed that reflection, differentiation, and assessment were important, but they indicated that these elements needed more explicit classroom guidance, structured prompts, and measurable indicators.

4.3 Qualitative Expert Feedback and Refinement

In addition to the FDM ratings, expert comments were analysed qualitatively. Expert comments confirmed the relevance of the model but suggested that several components needed clearer operationalisation. For example, the Differentiation required more explicit pathways for mixed-readiness classrooms. Reflection required structured prompts to avoid becoming a general closing activity. Assessment required indicators that capture both

products and processes. Debugging required teacher prompts that help students reason systematically rather than guess. These refinements are summarised in Table 4.

Table 4

Summary of Expert Qualitative Feedback and Refinement

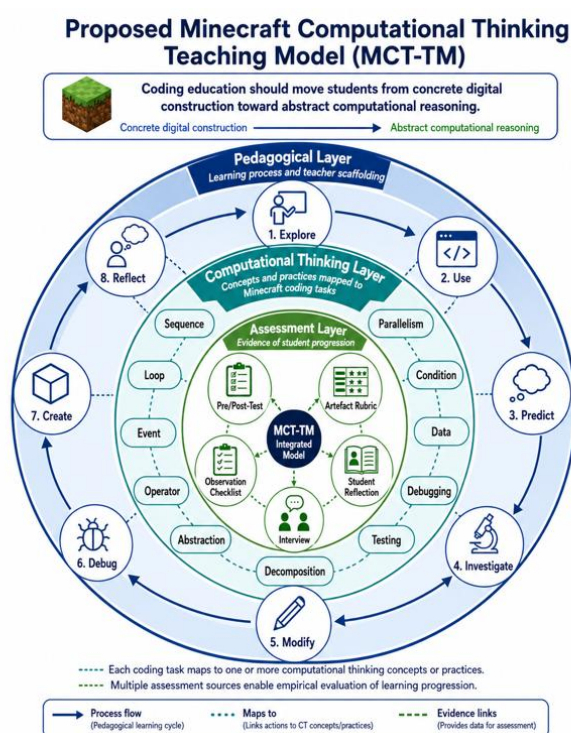
Area Refined	Expert Concern	Revision Made	Purpose
Differentiation	Students have different coding readiness.	Added beginner, intermediate, and advanced pathways.	Improves classroom practicality.
Reflection	Reflection may be too general.	Added prompts on concept used, error faced, debugging strategy, and improvement.	Improves CT explanation.
Assessment	Artefacts alone may not show CT.	Mapped evidence to tests, artefacts, observations, reflections, and interviews.	Improves assessment validity.
Debugging	Errors should be taught explicitly.	Added guided debugging questions and error-check routines.	Makes debugging a teachable CT practice.
Classroom implementation	Teachers need clearer lesson support.	Clarified teacher prompts, scaffolding, and evidence collection.	Improves implementation readiness.

4.4 Final Conceptual Model

The validated MCT-TM consists of three explicit relationships, as shown in Figure 2. First, the pedagogical cycle links the eight instructional stages as a repeated process of exploration, code use, prediction, investigation, modification, debugging, creation, and reflection. Second, each stage maps onto CT concepts and practices. Third, each activity is linked to assessment evidence so that learning is inferred from multiple sources rather than from completed artefacts alone.

Figure 2

Validated Minecraft Computational Thinking Teaching Model (MCT-TM)



5. Discussion

5.1 Positioning Minecraft as a Structured Coding Pedagogy

The findings extend existing research by repositioning Minecraft from being merely an engaging game-based learning environment to a pedagogically structured context for computational thinking (CT) development. Although previous studies have highlighted Minecraft's potential to promote engagement, creativity, collaboration, and various cognitive and academic outcomes (Slattery et al., 2025), such outcomes should not be conflated with the development of CT. A key limitation in literature is the tendency to assume that participation in building and coding activities naturally translates into computational understanding. Students may successfully construct complex artefacts, follow coding instructions, or complete challenges without fully understanding the computational principles that underpin their actions. This suggests that Minecraft itself is not the source of learning; rather, learning depends on how educators structure, scaffold, and assess students' interactions within the environment.

This interpretation aligns with the caution raised by Slattery et al. (2025), who noted that despite promising educational outcomes, the evidence base for Minecraft remains methodologically weak due to the prevalence of medium- and high-risk bias studies. Consequently, the MCT-TM should not be viewed as evidence that Minecraft inherently improves CT, but as a theoretically grounded response to a recognised pedagogical gap. Its primary contribution lies in making the learning process explicit rather than assuming that engagement will automatically produce understanding. By sequencing activities from exploration and guided use to prediction, investigation, modification, debugging, creation, and reflection, the model deliberately foregrounds the cognitive processes that underpin CT. More importantly, it acknowledges that conceptual understanding emerges through cycles of explanation, testing, error correction, and reflection rather than through construction alone. In doing so, the model shifts attention from the production of Minecraft artefacts to the development of computational reasoning, offering a more defensible framework for translating game-based participation into meaningful learning outcomes.

5.2 Supporting the Transition from Concrete Experience to Abstract Computational Reasoning

One of the key contributions of the MCT-TM is its explicit focus on supporting learners' progression from concrete Minecraft experiences to abstract computational understanding. This contribution is significant because novice programmers often struggle when computational concepts are introduced primarily through symbolic code detached from meaningful contexts. Minecraft reduces this barrier by providing visible and immediate representations of programming concepts, allowing learners to observe how sequences, loops, events, and debugging processes produce tangible changes in the game world. However, the educational value of these representations should not be overstated. While visual outcomes can increase accessibility and engagement, they do not automatically lead to conceptual understanding. Students may recognise patterns in game behaviour without fully understanding the computational principles that generate those outcomes.

The model responds to this challenge by treating concrete experience as an entry point rather than evidence of learning. Consistent with Bile's (2022) argument that Minecraft can make abstract computational ideas more accessible through interactive activity, the MCT-TM uses Minecraft challenges to establish meaningful contexts before gradually introducing code analysis, modification, debugging, and reflection. This progression also aligns with Mørch and Andersen's (2023) framework, which connects game-based building to increasingly abstract forms of programming. Importantly, their findings suggest that learners often experience

difficulty explaining the computational concepts embedded within functioning code, revealing a gap between successful task completion and genuine understanding. The MCT-TM addresses this limitation through the Predict, Investigate, Debug, and Reflect stages, which require students to justify predictions, explain code behaviour, analyse errors, and articulate computational reasoning. In this sense, the model shifts the focus from what students can build to what they can explain. Its contribution is therefore not merely the use of Minecraft as a learning tool, but the recognition that computational thinking develops when concrete experiences are deliberately transformed into opportunities for abstraction, explanation, and metacognitive reflection.

5.3 Addressing Uneven Development of Computational Thinking Concepts

The findings highlight the importance of explicitly scaffolding abstract computational thinking (CT) concepts rather than assuming that they will emerge naturally through Minecraft-based coding activities. Although Kutay and Öner (2022) reported improvements in students' understanding of computational concepts, their analysis revealed uneven learning across different CT components. Concepts such as sequences, events, loops, and parallelism were applied more successfully, whereas variables, operators, and conditionals remained comparatively difficult for students to use effectively. This suggests that the visual and interactive nature of Minecraft may make some computational concepts more accessible than others. Consequently, the successful completion of coding tasks should not be interpreted as evidence that learners have developed a balanced understanding of all CT concepts. Rather, certain abstract concepts may remain hidden beneath visible artefact construction and therefore require deliberate instructional support.

The MCT-TM addresses this challenge by embedding scaffolding stages that connect visible outcomes to underlying computational logic before students are expected to create original artefacts. Through the Investigate stage, learners examine relationships between code and programme output, while the Modify stage allows guided experimentation with parameters, coordinates, variables, and conditions. The Debug and Reflect stages further encourage students to explain discrepancies, justify decisions, and articulate the conceptual meaning of their code. These stages are particularly important for concepts such as variables, operators, and conditionals, which are not easily observable in the same way as loops or repeated structures. This finding suggests that Minecraft-based coding should not treat all CT concepts as equally learnable through exploration alone. Instead, effective instruction requires targeted pedagogical strategies that make less visible concepts explicit, discussable, and assessable. By requiring activities to be linked to specific CT concepts and practices, the MCT-TM moves beyond the assumption that engagement with Minecraft automatically leads to computational understanding and instead provides a more intentional pathway for conceptual development.

5.4 Debugging as a Central Learning Process

A notable finding from the expert validation was the strong support for debugging as a central component of the MCT-TM. This emphasis reflects a growing recognition in computational thinking literature that debugging is not merely a technical process of correcting errors but a cognitive process involving hypothesis formation, testing, evaluation, and revision. In Minecraft coding, debugging is particularly pedagogically valuable because programming errors often produce visible and immediate consequences in the game world. However, the visibility of errors alone does not guarantee learning. Students may identify and correct problems through trial-and-error approaches without understanding the underlying cause of the error. Consequently, treating debugging as an incidental activity risks overlooking its potential to develop deeper computational reasoning.

The MCT-TM addresses this limitation by positioning debugging as an explicit instructional stage rather than a consequence of programming mistakes. This approach aligns with the Action–Breakdown–Repair perspective, which views breakdowns as opportunities for conceptual development rather than evidence of failure (Mørch & Andersen, 2023). Similarly, Kutay and Öner (2022) found that testing and debugging are among the most frequently applied computational thinking practices in Minecraft coding activities. Nevertheless, the educational benefits of debugging depend on appropriate scaffolding. Without structured guidance, learners may focus on fixing errors efficiently rather than understanding why those errors occurred. By encouraging students to predict expected outcomes, analyse discrepancies, justify revisions, and evaluate solutions, the MCT-TM transforms debugging into a reflective and systematic process. Its contribution, therefore, lies not simply in recognising the importance of debugging, but in making the reasoning embedded within debugging visible, accessible, and central to computational thinking development.

5.5 Applicability in Southeast Asian Educational Contexts

MCT-TM has potential relevance for Southeast Asian primary education because many systems in the region are increasingly integrating coding, digital competence and computational thinking into curriculum reform. However, applicability depends on three practical conditions. First, curriculum alignment is needed: Minecraft tasks should be mapped to national learning standards, topics, and assessment expectations rather than treated as enrichment only. Second, teacher readiness is essential: teachers need examples of prompts, debugging routines, classroom management strategies, and assessment rubrics. Third, resource availability must be considered: schools differ in device access, internet stability, Minecraft Education licensing, language proficiency and pupils' prior experience with coding. These considerations suggest that MCT-TM should be implemented flexibly. In well-resourced schools, students may progress from block-based coding to Python or JavaScript extensions. In lower-resource contexts, teachers may use shared devices, pair programming, unplugged planning, screenshots, code cards or teacher-led demonstrations before independent practice. The model's differentiated pathways are therefore important for Southeast Asian classrooms where learner readiness and school infrastructure may vary considerably.

5.6 Implications

For teachers, MCT-TM offers a lesson structure that prevents students from moving too quickly into unstructured buildings. For curriculum designers, it provides a way to align Minecraft tasks with CT concepts, practices, and evidence. For researchers, it offers a validated model that can be tested using quasi-experimental or mixed-methods designs. For assessment, the model encourages triangulation: tests can capture conceptual knowledge, artifact rubrics can capture functional and creative products, observations can capture collaboration and debugging behaviour, and reflections or interviews can capture students' computational explanations.

5.7 Limitations and Future Research

This study is limited to model development and expert validation. It does not include classroom implementation, student performance data, or teacher enactment data. The expert panel was purposely selected and relatively small, which is appropriate for FDM but not a substitute for implementation evidence. Future research should test MCT-TM through pilot lessons, teacher feedback, student artefact analysis, and pre/post CT measures. Stronger studies may include comparison groups, delayed post-tests, and qualitative analysis of debugging logs, reflection journals, and classroom interaction. Attention should be given to difficult concepts such as variables, operators, and conditionals.

6. Conclusion

In conclusion, the MCT-TM offers a structured approach for teaching coding through Minecraft. Its main strength lies in its integration of playful exploration, code comprehension, modification, debugging, artefact creation, and reflection. The model responds to limitations in existing Minecraft coding studies by providing a clear pedagogical sequence, explicit computational thinking mapping, and multiple forms of assessment. While further empirical testing is required, the proposed model has strong potential to guide teachers, curriculum designers, and researchers in transforming Minecraft from an engaging digital game into a purposeful computational thinking pedagogy for coding education.

Acknowledgment

The author would like to express their sincere appreciation to the experts for their voluntary participation in this study. Their insights and willingness to share their professional experiences made this research possible. The limited use of generative AI (OpenAI's ChatGPT) supported language clarity and formatting. All AI-assisted outputs were critically reviewed and validated by the author, who assumes full responsibility for the content.

References

- Bile, A. (2022). Development of intellectual and scientific abilities through game-programming in Minecraft. *Education and Information Technologies*, 27(5), 7241–7256. <https://doi.org/10.1007/s10639-022-10894-z>
- Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American Educational Research Association* (pp. 1–25). American Educational Research Association. <http://scratched.gse.harvard.edu/ct/files/AERA2012.pdf>
- Giannakoulas, A., & Xinogalos, S. (2024). A critical review of primary school students' difficulties in learning programming through educational games. *Journal of Educational Computing Research*, 63(3), 627–661. <https://doi.org/10.1177/07356331241309074>
- Kutay, E., & Öner, D. (2022). Coding with Minecraft: The development of middle school students' computational thinking. *ACM Transactions on Computing Education*, 22(2), Article 21, 1–19. <https://doi.org/10.1145/3471573>
- Mørch, A. I., & Andersen, R. (2023). Programming with Minecraft Bedrock Up: Modeling, coding, and computational concepts. In L. D. Spano, A. Schmidt, C. Santoro, & S. Stumpf (Eds.), *End-user development: 10th International Symposium on End-User Development, IS-EUD 2023* (pp. 230–240). Springer. https://doi.org/10.1007/978-3-031-34433-6_15
- Nkadimeng, M., & Ankiewicz, P. (2022). The affordances of Minecraft education as a game-based learning tool for atomic structure in junior high school science education. *Journal of Science Education and Technology*, 31(5), 605–620. <https://doi.org/10.1007/s10956-022-09981-0>
- Van Schoors, R., Bhatt, S., Elen, J., Raes, A., Van Den Noortgate, W., & Depaepe, F. (2024). Design and development of a digital personalized learning track: Bridging the gap between textual and visual programming. *International Journal of Designs for Learning*, 15(1), 74–95. <https://doi.org/10.14434/ijdl.v15i1.35224>

- Slattery, E. J., Lehane, P., Butler, D., O'Leary, M., & Marshall, K. (2025). Assessing the benefits of digital game-based learning with Minecraft in children, adolescents and young adults: A broad systematic review. *Review of Education*, 13(1), Article e70035. <https://doi.org/10.1002/rev3.70035>
- Sentance, S., & Waite, J. (2017). PRIMM: Exploring pedagogical approaches for teaching text-based programming in school. In *Proceedings of the 12th Workshop on Primary and Secondary Computing Education* (pp. 113–114). Association for Computing Machinery. <https://doi.org/10.1145/3137065.3137084>
- Suwannik, W. (2025). Minecraft: An engaging platform to learn programming. In *2025 IEEE/ACIS 23rd International Conference on Software Engineering Research, Management and Applications (SERA)* (pp. 543–546). IEEE. <https://doi.org/10.1109/SERA65747.2025.11154595>
- Voštinár, P., & Dobrota, R. (2022). Minecraft as a tool for teaching online programming. In *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)* (pp. 648–653). IEEE. <https://doi.org/10.23919/MIPRO55190.2022.980338>

STUDENT

SECTION

EFFICIENCY ASSESSMENT OF A FOOD WASTE SHREDDING MACHINE FOR BIOGAS PRODUCTION

Marie Antonette Anahaw
Ashly Claudet Grafilo
Harvey Oborza*
Cris Edison U Villanueva

Lutucan Integrated National High School, Philippines
<hperezoborza@gmail.com>

ABSTRACT

Food waste is a growing environmental concern that requires practical and sustainable waste management solutions, particularly in schools and local communities. This study evaluated the performance of a newly developed food waste shredding machine and was investigated to its fullest potential as a mechanical pretreatment method for improving food waste utilisation and biogas production through anaerobic digestion. The machine was tested using common kitchen waste, including rice, vegetables, and mixed food waste, to determine its shredding efficiency, processing rate, and ability to reduce particle size. Results showed that the machine consistently reduced food waste particles to less than one centimetre, demonstrating effective shredding performance. Among the tested materials, rice achieved the highest shredding efficiency of 98.8% and the fastest processing rate of 2.47 kg/min. The reduction in particle size also improved the machine's anaerobic digestion process, producing greater biogas output and a more stable and flammable flame compared with unprocessed food waste. Statistical analysis using one-way analysis of variance (ANOVA) showed no significant difference in shredding efficiency among the three types of food waste, $F(2,6) = 3.01$, $p = 0.12 > 0.05$. This indicates that although variations in performance were observed, the differences were not statistically significant. Overall, the findings demonstrate that mechanical shredding can serve as an effective pretreatment method for food waste utilisation and biogas generation. The developed machine provides a practical, low-cost, and sustainable approach to small-scale waste management while supporting waste reduction and renewable energy production.

Keywords: Food Waste, Shredding Machine, Anaerobic Digestion, Biogas Production, Shredding Efficiency, Waste Management

1. Introduction

Food waste is one of the most pressing global issues: approximately one-third of all food produced worldwide, or about 1.3 billion tons annually, is wasted. Much of this waste decomposes anaerobically in landfills and produces methane (CH_4), a greenhouse gas 25 times more potent than carbon dioxide (CO_2) (Fatemi et al., 2024; Lewis, 2022). According to Lai (2021), this makes food waste the third-largest contributor to global greenhouse gas emissions. It also represents a significant loss of valuable resources such as water, energy, and labour.

In the Philippines, the problem is evident in households, markets, and especially schools, where canteens generate organic waste daily. Improper disposal attracts pests, produces foul odours, and compromises sanitation. More importantly, unmanaged food waste limits opportunities for resource recovery and renewable energy production (Majores, 2015).

Anaerobic digestion (AD) has emerged as a viable solution for managing food waste. It produces biogas, a renewable mixture of methane and carbon dioxide, while simultaneously reducing organic pollution (Milano et al., 2016; Jameel et al., 2024; Anderson, 2025). Studies emphasise that AD efficiency depends heavily on feedstock characteristics, particularly particle size (Dölle et al., 2020). Pre-treatment through shredding accelerates microbial decomposition, increases surface area, and enhances gas yield (Shan, 2024). According to Li et al. (2011), shredding also improves waste management by reducing bulk, facilitating storage, and producing uniform material suitable for composting and digestion.

However, while most studies focus on biogas yield from large industrial digesters, there is limited data on the engineering performance of shredders themselves. Specifically, few studies evaluate shredding efficiency, throughput rate, and particle size reduction in small-scale, low-cost machines for school or community settings. Evaluating these engineering parameters is critical to determining whether such machines are feasible for everyday applications (Galvez et al., 2023).

Therefore, this study evaluates the performance of a food waste shredding machine designed for use in schools and communities. Specifically, it investigates its shredding efficiency, throughput rate, and particle size reduction across three types of waste: rice, vegetables, and mixed canteen scraps. A secondary validation through anaerobic digestion was also conducted to determine whether shredding improved biogas yield and methane content. In school canteens, where organic waste is generated daily but space and budget are limited, an efficient shredder could make anaerobic digestion a practical waste-to-energy option. By focusing on engineering performance, this study contributes to sustainable food waste management strategies and supports renewable energy initiatives in educational and community settings (Mertins & Wawer, 2022).

2. Methodology

2.1 Research Design

A true experimental research design was used to assess the operational performance of the food waste shredder prototype. The independent variable was the type of food waste, categorised into three treatments: rice, vegetables, and mixed scraps. The dependent variables measured to evaluate performance included: shredding efficiency (%), throughput rate (g/min), and post-shredding particle size (cm). To ensure statistical reliability, each treatment was replicated three times, resulting in a total of nine experimental trials ($n=9$).

2.2 Materials

The fabrication of the food waste shredder prototype involved a structural metal frame, high-grade stainless-steel cutting blades, an electric motor, an intake hopper, a mechanical safety guard, and a designated collection bin. Food waste samples were collected directly from the school canteen to mimic real waste management conditions. Data collection used the following calibrated instruments: a digital scale for precise mass measurements, a digital stopwatch to track operational time, a ruler and sieve mesh to measure particle sizes, and an airtight anaerobic digester to evaluate biogas potential from the shredded material.

2.3 Procedure

Fresh food waste samples from the school canteen were immediately sorted into three categories: rice, vegetables, and mixed scraps. To control for mass variations, each sample was standardised to exactly 1.0 kg (1,000 g) before processing. Each batch was fed uniformly

into the intake hopper, and the machine was activated immediately afterward. Processing time was recorded with a digital stopwatch, starting when the substrate engaged the blades and ending when all material was collected. The processed output was weighed to determine mass recovery and randomly sampled for particle size analysis.

- Shredding Efficiency was calculated using the formula:

$$\text{Efficiency} = \frac{\text{output weight}}{\text{input weight}} \times 100$$

- Throughput was calculated:

$$\text{Throughput} = \frac{\text{output weight}}{\text{processing time}} \times 100$$

- Particle Size: Five random samples per batch were measured to compute the average.

For biogas validation, 500 g of shredded and unshredded food waste were placed into identical airtight digesters. Over three days, the researchers observed gas production and tested its flammability.

2.4 Data Analysis

The data gathered from the shredding machine trials were systematically tabulated and analysed to evaluate its performance across three food waste categories: rice, vegetables, and input weight multiplied by 100. Throughput rate (g/min) was determined by dividing the output mass by the processing time, while particle size reduction was obtained by averaging measurements from randomly selected shredded samples.

For statistical analysis, a one-way Analysis of Variance (ANOVA) was performed at the 0.05 level of significance to test for differences in shredding efficiency among the three waste categories.

To test whether differences among the three treatments were statistically significant, a one-way Analysis of Variance (ANOVA) was performed at an alpha level of 0.05. Descriptive statistics (mean and standard deviation) and inferential statistics (F -statistic and p -value) were computed using Microsoft Excel 2019. The software handled data tabulation, formula verification, and automatic ANOVA generation, ensuring accurate calculations and reducing human error.

Figure 1 provides a visual representation and step-by-step documentation of the research process, starting with the construction of the machine itself. The flow then transitions to the setup for the anaerobic digestion validation, which includes attaching the biodigester to the machine and securing the biodigester components, such as attaching the hose and attaching a pipe for the mixer. A critical step documented in the flow is making sure the biodigester is oxygen-free to ensure anaerobic conditions are met for optimal biogas production. The process concludes with attaching an inflatable rubber for biogas storage, leading to the tentative product for flammability testing. This flowchart ensures that all steps, both mechanical and biological, are clearly documented.

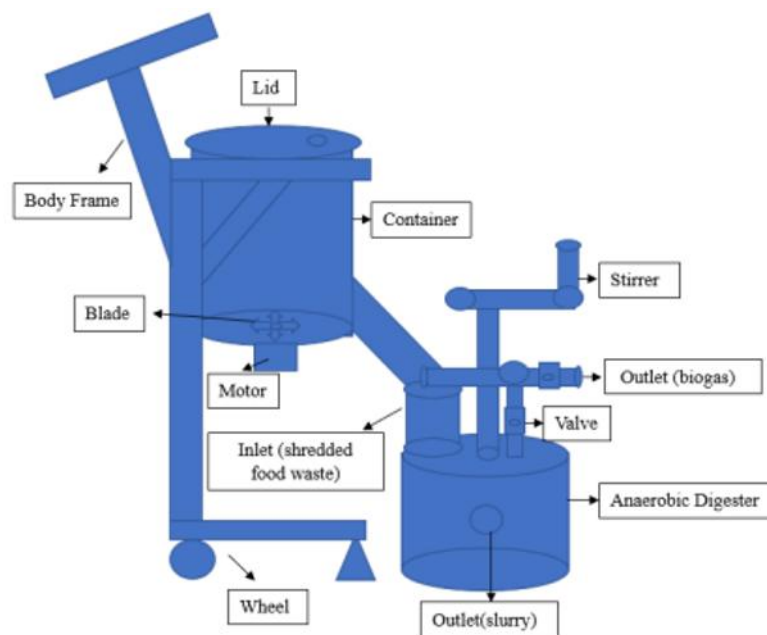
Figure 1*Flowchart of the Procedure***Figure 2***Schematic Diagram*

Figure 2 presents the schematic diagram of the custom-built food waste shredding machine. The diagram identifies all key components, including the stainless-steel blades, the hopper where the food waste is fed, the motor that drives the shredding mechanism, and the collection bin. This visual aid is essential for understanding the mechanical components referenced in the methodology and providing context for the engineering performance results.

3. Result and Discussion

Shredding Efficiency

Rice consistently showed the highest shredding efficiency (98.8%), followed by mixed waste (98.3%), while vegetables recorded the lowest (96.6%). The fibrous nature of vegetables explains their lower efficiency compared to the softer texture of rice.

Table 1

Shredding Efficiency by Food Waste Types

Food Waste Type	Trial 1	Trial 2	Trial 3	Mean (%)
Rice	98	98.8	99.6	98.8
Vegetables	94.6	98.2	97.2	96.6
Mixed Waste	98.8	99	98.2	98.3

The shredding efficiency of the machine ranged from 96.6% to 98.8%, demonstrating that the prototype successfully minimised waste loss during processing. Among the three waste categories, rice recorded the highest mean shredding efficiency (98.8%), followed by mixed scraps (98.3%) and vegetables (96.6%). The superior performance with rice can be attributed to its uniform grain structure and lower fibre content, which made it easier for the blades to cut consistently. Conversely, vegetables, with their fibrous composition and variable moisture levels, posed greater resistance, slightly lowering efficiency.

Table 2

Average Throughput Rate

Food Waste Type	Average Throughput Rate (g/min)
Rice	2.47
Vegetables	2.41
Mixed Waste	2.45

The throughput rate provides insight into the speed and capacity of the shredding machine. The machine processed rice at 2.47 g/min, which was faster than vegetables (2.41 g/min) and mixed waste (2.45 g/min). The higher throughput with rice is consistent with its high shredding efficiency, suggesting that uniform texture not only improves conversion but also enhances processing speed. Vegetables, on the other hand, slowed down the machine. Their high fibre content and bulkier structure required more blade resistance, thereby extending processing time. Mixed waste demonstrated intermediate throughput, reflecting its balance of soft and fibrous components.

Table 3

Mean Particle Size

Food Waste Type	Mean Particle Size
Rice	0.77
Vegetables	3.77
Mixed Wastes	1.5

One of the primary goals of shredding is to reduce particle size, thereby increasing surface area for microbial decomposition during anaerobic digestion. The results showed that the shredder successfully reduced all waste categories to less than 5 cm, which is the threshold recommended for efficient anaerobic digestion.

Rice again achieved the smallest particle size (0.77 cm), consistent with its structure, while vegetables remained relatively larger (3.77 cm). Mixed waste averaged 0.58 cm, falling between the two categories. These results confirm that the machine is capable of producing a uniform, manageable feedstock suitable for both composting and biogas production.

Table 4

One-way ANOVA Results for Shredding Efficiency

Source of Variation	SS	df	MS	F	p-value	Decision ($\alpha = 0.05$)
Between Groups	8.57	2	7.79	4.28	3.01	Accept H0
Within Groups	8.53	6	0.36	1.42		
Total	17.10	8				

The one-way ANOVA revealed that statistically, there is no significant difference in shredding efficiency across the three waste categories, $F(2,6) = 4.28$, $p = 3.01 < 0.05$. This result indicates that the type of food waste does not directly affect machine efficiency.

From an engineering perspective, the implication is clear: while the current blade design works effectively for soft-textured wastes, future iterations of the machine may require blade adjustments or motor upgrades to optimise processing of fibrous wastes such as vegetables.

Acknowledgement

The researchers would like to extend their deepest gratitude to all individuals who contributed to the successful completion and final presentation of this study.

First and foremost, our profound appreciation goes to Mr Cris Edison U Villanueva, our Grade 10 Research Adviser, for his essential guidance, meticulous review, and unwavering support in shaping and refining the final written paper. His expertise was vital in ensuring the clarity and scholarly quality of the manuscript for submission. We would also express our sincere thanks to Ms Joey Lyne Datinggaling, our Grade 9 Research Adviser, for her foundational support, insightful direction, and encouragement during the initial conceptualisation and early stages of this project.

We wish to sincerely acknowledge the technical and academic assistance provided by others. We thank Mr Obet Tantay and Mr Ariel Magpantay for their valuable help in conceptualising, visualising, and refining the design of the shredding machine, and for sharing tips on the practical application of food waste for biogas production. Our heartfelt thanks likewise go to Mr Daniel Ibarrola, whose guidance in proper data-gathering methods and computation techniques ensured the accuracy and reliability of our quantitative results. We are equally grateful to Mr Israel Oborza for his practical support and hands-on involvement in constructing the machine used in this study, which brought our design into reality.

Lastly, we extend our warmest appreciation to our classmates, whose constant encouragement, cooperation, and moral support motivated us throughout the entire research process. Their enthusiasm and belief in our capabilities made this journey meaningful and

fulfilling. We have immense gratitude for every single person who contributed in any way, whether it was through support or knowledge. We extend our sincerest gratitude to those who have helped us succeed and supported us along our journey.

References

- Anderson, K. (2025, June 03). What is Biogas and is it sustainable? Greenly Earth. <https://greenly.earth/en-gb/blog/company-guide/what-is-biogas-and-is-it-sustainable>
- Badurek, C.A. (2026, July 10). Biogas. Encyclopedia Britannica. <https://www.britannica.com/technology/biogas>
- Dölle, K., Hughes, T., & Kurzmann, D. (2020). From fossil fuels to renewable biogas production from biomass-based feedstock: A review of anaerobic digester system. *Journal of Energy Research and Reviews*, 5(3), 1–37. <https://doi.org/10.9734/jenrr/2020/v5i330147>
- Fatemi, S. F., Eini-Zinab, H., Anari, F. M., Amirolad, M., Babaei, Z., & Sobhani, S. R. (2024). Food waste reduction and its environmental consequences: A quasi-experimental study in a campus canteen. *Agriculture & Food Security*, 13, Article 37. <https://doi.org/10.1186/s40066-024-00488-y>
- Galvez, J. O., De Silos, P. Y., Paguntalan, A. M., Reyes, J. K., & Pedrigal, Z. (2023). Assessing the performance of biogas digester as a sustainable energy technology: Solution for domestic use in Tunasan, Muntinlupa City, Philippines. *Physical Science & Biophysics Journal*, 7(1), Article 000251. <https://doi.org/10.23880/psbj-16000251>
- Lai, P. (2021, November 22). *Explainer: What is food waste?* Earth.Org. <https://earth.org/what-is-food-waste/>
- Lewis, J. (2022, October 17). How does Food Waste Affect the Environment? Earth.Org. <https://earth.org/how-does-food-waste-affect-the-environment/>
- Li, Y., Park, S. Y., & Zhu, J. (2011). Solid-state anaerobic digestion for methane production from organic waste. *Renewable and Sustainable Energy Reviews*, 15(1), 821–826. <https://doi.org/10.1016/j.rser.2010.07.042>
- Mertins, A., & Wawer, T. (2022). How to use biogas? A systematic review of biogas utilization pathways and business models. *Bioresources and Bioprocessing*, 9, Article 59. <https://doi.org/10.1186/s40643-022-00545-z>
- Milano, J., Ong, H. C., Masjuki, H. H., Chong, W. T., Lam, M. K., Loh, P. K., & Vellayan, V. (2016). Microalgae biofuels as an alternative to fossil fuel for power generation. *Renewable and Sustainable Energy Reviews*, 58, 180–197. <https://doi.org/10.1016/j.rser.2015.12.150>
- Jameel, M. K., Mustafa, M. A., Ahmed, H. S., Mohammed, A. J., Ghazy, H., Shakir, M. N., & Lawas, S. K. M. (2024). Biogas: Production, properties, applications, economic and challenges: A review. *Results in Chemistry*, 7, Article 101549. <https://doi.org/10.1016/j.rechem.2024.101549>
- Majores, J. G. (2015). Implementing biogas technology project in Malvar, Batangas, Philippines. *Asia Pacific Journal of Multidisciplinary Research*, 3(4), 158–164. https://www.researchgate.net/publication/341051636_Implementing_Biogas_Technology_Project_in_Malvar_Batangas_Philippines
- Shan. (2024, March 14). *Shredder optimization for biogas production from food waste*. GEP ECOTECH. <https://www.gepecotech.com/release/blog/shredder-optimization-for-biogas-production-from-food-waste.html>

DEVELOPMENT AND PERFORMANCE EVALUATION OF AN AI-INTEGRATED ULTRASONIC FLY REPELLENT SYSTEM WITH REAL-TIME HOUSE FLY DETECTION

Kanyanat Khamhongsa*
Premyuda Witchayateerakun
Thanaphon Phunamsiang
Chumpon Chareesaen

Kalasinpittayasan School, Kalasin, Thailand
<bamee9255@gmail.com>

ABSTRACT

This study aimed to develop and evaluate an ultrasonic fly repellent system integrated with an artificial intelligence-based house fly detection system. This research project, conducted by Ms Kanyanat Khamhongsa, Ms Premyuda Witchayateerakun, and Ms Thanaphon Phunamsiang under the supervision of Mr Chumpon Chareesaen at Kalasinpittayasan School, Thailand, sought to provide a sustainable, chemical-free alternative to traditional pest control. The research comprised two primary components: investigating the influence of ultrasonic frequency and emitter distance on house fly landings, and developing an automated real-time detection system utilising the YOLO11 object detection model. Experimental results demonstrated that increasing the ultrasonic frequency from 20 to 35 kHz substantially reduced the average number of house flies landing on food sources from 13.00 ± 1.73 to 4.33 ± 1.15 individuals, with peak effectiveness at 35 kHz. Furthermore, the ultrasonic emitter performed optimally when positioned within 0–5 cm of the food source. Regarding the AI component, the YOLO11n model achieved a precision of 0.774, recall of 0.889, an F1-score of 0.827, and an mAP50 of 0.832, with an average inference speed of 6.7 milliseconds per image. Integrated system testing demonstrated a 95% speaker activation accuracy with a 1–2 second response time. This project provided invaluable practical experience in designing IoT hardware, training deep learning models, and managing complex system workflows. The AI-integrated ultrasonic system has demonstrated significant effectiveness in reducing house fly populations. It is practical for real-world application and possesses strong potential to serve as an innovative, safe, and sustainable replacement for chemical insecticides in households and food-preparation environments. Future developments will focus on expanding the operational coverage area, enhancing the working distance, and refining the design for commercial-scale production.

Keywords: Artificial Intelligence, House Fly, Ultrasonic Repellent System, YOLO11, Real Time Detection

1. Introduction

The house fly (*Musca domestica* Linnaeus, Diptera: Muscidae) is one of the most common insect species worldwide and is closely associated with human habitats, particularly in areas containing food waste, organic refuse, and human or animal excreta. House flies are recognised as mechanical vectors capable of transmitting more than 100 pathogenic microorganisms responsible for several diseases, including diarrhoea, typhoid fever, cholera, dysentery, and certain skin infections. These diseases pose significant threats to public health and sanitation (Department of Disease Control, 2021). The biological and behavioural characteristics of house flies, such as feeding on food residues, garbage, and faecal matter,

facilitate the attachment of microorganisms to their legs, body surfaces, and mouthparts. When flies subsequently land on food, food preparation surfaces, or utensils used by humans, pathogens may be transmitted through saliva, faeces, eggs, or larvae, resulting in food contamination and an increased risk of gastrointestinal diseases (Khamesipour et al., 2018). Currently, chemical insecticides remain the primary method for controlling house fly populations because of their rapid effectiveness and ease of application. However, the continuous and excessive use of chemical pesticides has resulted in the accumulation of toxic residues in the environment, food chains, and human bodies. Furthermore, prolonged exposure to insecticides has contributed to the development of pesticide resistance, enabling insects to survive previously effective concentrations of chemical agents. Consequently, higher doses or alternative insecticides are often required, leading to increased economic costs and environmental burdens (Liu et al., 2019).

Previous studies have demonstrated that house flies are capable of detecting ultrasonic waves through vibration-sensitive sensory organs that respond to fluctuations in air pressure. Frequencies ranging from approximately 20 to 40 kHz have been reported to interfere with the sensory, navigation, and balance mechanisms of flies, thereby inducing avoidance behaviour when exposed to ultrasonic sound (Sari & Wijaya, 2020). Moreover, ultrasonic waves can create unfavourable environmental conditions for flies, reducing their tendency to approach food sources and inhabit affected areas without introducing chemical contamination into food products or the surrounding environment (Wang et al., 2022).

Despite these findings, most existing ultrasonic repellent systems operate continuously without intelligent activation, resulting in unnecessary energy consumption and reduced operational efficiency. In addition, limited research has integrated real-time AI-based fly detection with ultrasonic repellent technology. Consequently, there remains a need for a system that can detect house flies in real time and activate the ultrasonic repellent only when flies are present. Such an approach could improve energy efficiency, enhance practical applicability, and increase the effectiveness of fly repellent systems in real-world environments.

Therefore, this study proposes an AI-integrated ultrasonic fly repellent system equipped with a real-time house fly detection mechanism based on the YOLO11 object detection model. The system is designed to automatically detect house flies and activate the ultrasonic repellent only when required. By combining AI with ultrasonic repellent technology, this study aims to develop and evaluate a more intelligent, energy-efficient, and environmentally friendly fly repellent solution for household and food-preparation environments.

2. Literature Review

2.1 House Fly (*Musca domestica*)

The house fly (scientific name as *Musca domestica*) is an insect belonging to the order Diptera, commonly found around residential areas. The house fly undergoes holometabolous metamorphosis (complete metamorphosis), consisting of 4 stages: egg, larva, pupa, and adult. The entire life cycle takes approximately 7–14 days, depending on temperature and environmental conditions. House flies prefer to inhabit and lay eggs in areas with food scraps, organic waste, or animal manure, which serve as food sources for the larvae, with adult flies having an average lifespan of 2–4 weeks.

Flies present widespread problems for humans regarding health, environment, living conditions, and economics, particularly in communities with accumulated waste, food remnants, or manure. House flies have a high reproductive rate and thrive in warm, humid

conditions, allowing their population to increase rapidly within a short period. Their lifestyle, which relies on decaying organic matter for sustenance and oviposition, turns them into critical vectors that directly transport pathogens from unsanitary sources into human dwellings or food. Transmission occurs when flies use their mouthparts and legs to touch animal faeces, sewage, or garbage, and then land on food or items used by humans. Pathogens are subsequently deposited through regurgitation and defecation, which are normal fly behaviours. Furthermore, the hairs on their legs and bodies can easily trap microbes, causing contamination without requiring any biting or blood-sucking. Consequently, gastrointestinal diseases such as cholera, typhoid, dysentery, and food poisoning can spread rapidly via flies acting as vectors.

2.2 Fly Perception and Behavioural Responses to Sound Frequencies

Flies can perceive high-frequency waves beyond human hearing thresholds through an organ called Johnston's organ, located at the base of the antennae (Eberl & Boekhoff-Falk, 2002). This organ detects air movement caused by sound waves. When high-frequency waves hit the antennae, they induce vibrations, which Johnston's organ converts into neural signals sent to the brain, allowing the fly to perceive the sound waves. Flies can perceive sound vibrations at relatively low to moderate frequencies; low-to-mid frequency sounds like wind, animal movements, or environmental noise often trigger heightened defensive behaviours, as these sounds may signify predators or disturbances. Meanwhile, specific high-frequency ranges can disrupt the nervous system and reduce flying activities. In some cases, flies will avoid areas with continuous loud noises or vibrations because such factors prevent them from efficiently conducting activities like foraging or mating.

Additionally, research indicates that certain extreme high-frequency sound waves can impact the central nervous system of flies, slowing down their movements or causing physiological stress. Certain sound levels can disrupt their balance because fly wings rely on the ambient air vibration frequency to stabilise flight. If high-frequency sound waves interfere with the surrounding air vibrations, flight directional control becomes impaired (Albert & Göpfert, 2013).

2.3 Frequency Ranges Affecting Flies

Regarding fly behaviour toward sound waves, studies indicate that the frequency range with the most statistically significant impact on flies is ultrasonic waves between 25–35 kHz, which is the most effective range for repelling flies (Raman & Kumar, 2017). Waves within this frequency disrupt the sensory systems and acoustic organs of flies, causing irritation and behavioural alterations: flies exhibit annoyance, restlessness, and a tendency to fly away from the wave source. Continuous ultrasonic emissions disrupt the navigation and balancing systems of flies, rendering their flight unstable (Loh & Wilson, 2015).

2.4 Human Impacts of Ultrasonic Waves

Ultrasonic waves used in medical and industrial applications today are considered safe when operated under appropriate supervision (Chen et al., 2026). However, exposure to high-intensity ultrasonic waves or improper environmental settings can affect humans in several ways, such as impacts on the auditory system. Although ultrasonic waves fall outside the human hearing range, continuous exposure to high-intensity waves can cause adverse symptoms like tinnitus, dizziness, or ear discomfort due to internal ear structures experiencing vibrations (Chen et al., 2026). However, these effects typically occur in industrial environments utilising high-frequency machinery with wave intensities significantly higher than household appliances.

2.5 Impacts of Ultrasonic Waves on Other Animals

Biologically, cats possess a broad hearing range, perceiving frequencies from approximately 45 Hz up to 85 kHz (Heffner, R.S., & Heffner, H.E., 1985), or at least 64 kHz. This indicates that cats have a high capacity to perceive ultrasonic waves exceeding 20 kHz and have the potential to experience annoyance or disturbance from them. Meanwhile, dogs have a hearing range of approximately 67 Hz to 45 kHz. Therefore, ultrasonic waves in the 20 kHz to 45 kHz range directly affect dogs, which is the same frequency band emitted by dog training whistles and typical animal repellents (LSU). The primary impact of ultrasonic waves on dogs and cats manifests as behavioural modifications, indicating discomfort or avoidance efforts, particularly when the sound pressure level (SPL) is high. Research testing ultrasonic cat repellents emitting 19 kHz to 30 kHz waves confirmed that these intensity levels are sufficient to prompt negative responses. In cats, behaviours observed under high-intensity ultrasonic waves include significant signs of disturbance, such as frequent ear flicking, decreased exploratory activities, and attempts to avoid the wave source area. Dogs exhibit more visible signs of stress and anxiety, such as restlessness, anxiety, cowering, or attempts to leave environments with high-intensity ultrasonic waves. These behaviours reflect that high-intensity ultrasonic waves can cause discomfort and act as stressors for high-frequency-sensitive mammals (Wag! Editorial Team, 2022).

3. Methodology

3.1 The Experiment

Experiment 1: Determination of the Optimal Ultrasonic Frequency for House Fly Repellence

- 1) An experimental cage was designed and constructed to house and confine 20 houseflies during the test, which was conducted in a closed, temperature- and humidity-controlled room to eliminate external environmental interference.
- 2) The ultrasonic transducer system was configured to generate ultrasonic frequencies of 20, 25, 30, and 35 kHz.
- 3) A video recording system was installed to monitor and record the behavioural responses of house flies following exposure to each ultrasonic frequency.
- 4) This experiment was conducted with three consecutive replicates.

Figure 1

Shows the Fly Rearing and Containment Equipment



Experiment 2: Investigation of the Effect of Distance Between the Ultrasonic Transducer and the Food Source on Fly Repellence Efficiency

- 1) The completed ultrasonic transducer device was positioned at distances of 0 cm, 5 cm, 10 cm, and 15 cm from the fly feeding container.
- 2) The ultrasonic generator was activated at the selected frequency (35 kHz) for 5 minutes.
- 3) The behaviours of house flies were observed and recorded, including escape responses, aggregation behaviour, movement away from the food source, and continued landing on the food source.
- 4) Each experiment was repeated at least three times for every frequency tested. The number of flies before and after activation of the device was recorded, and the time required for flies to leave the food source was measured to evaluate the effectiveness of the ultrasonic fly repellent system.

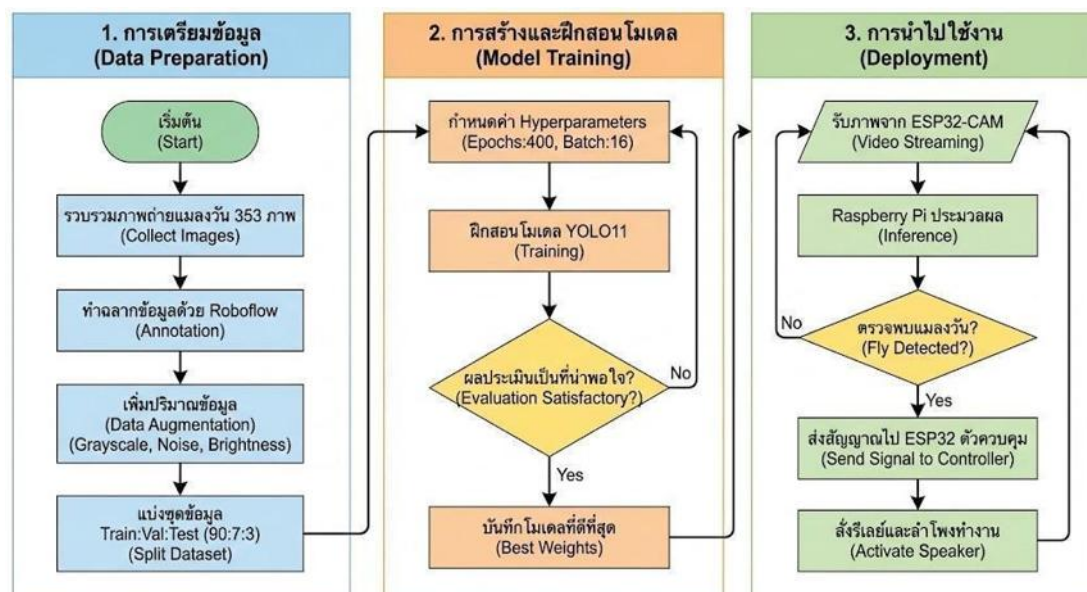
3.2 The Development and Performance Evaluation of an Artificial Intelligence-Based Fly Detection System Integrated with the Ultrasonic Fly Repellent Device

3.2.1 The System Overview

In developing the proposed system, the researchers designed a comprehensive workflow integrating computer vision, artificial intelligence, and ultrasonic fly control technologies. The system was developed to automatically detect the presence of house flies and activate the ultrasonic repellent device only when flies were identified, thereby improving operational efficiency and reducing unnecessary energy consumption. The system workflow consisted of three main stages: Data Preparation, Model Training, and Deployment.

Figure 2

System Workflow of the Proposed System



3.2.2 Tools and Equipment

1) Hardware

- Training Server
- CPU: Intel® Core™ i7-14700HX Processor
- GPU: NVIDIA® GeForce RTX™ 5060 Laptop GPU with 8 GB GDDR7
- RAM: 16 GB DDR5-4800 MT/s (SODIMM)
- Deployment Hardware
- ESP32-CAM: Used for real-time image acquisition and video streaming.
- Raspberry Pi: Used for real-time image processing and object detection inference using the trained YOLO11 model.

2) Software

- Programming Language: Python 3.11.9
- Main Libraries: PyTorch and Ultralytics YOLO Framework
- Data Preparation Tool: Roboflow
- Visual Studio Code: Used for model training
- Programming Languages: Python and C++

3) Data Preparation

- a) **Data Collection** The researchers collected images of house flies from various sources, including self-captured images and publicly available datasets such as Kaggle. A total of 353 house fly images were collected, covering a wide range of environmental conditions, including bright lighting, low-light conditions, and different viewing angles.
- b) **Data Annotation** Object annotation was performed using the Roboflow platform. Bounding boxes were manually drawn around each target object, and all objects were assigned to a single class, namely Fly.
- c) **Dataset Splitting** The dataset was divided into three subsets for model training and evaluation, as shown in Table 1.

Table 1

Dataset Splitting for Model Training and Evaluation

Dataset	Number of Images	Percentage	Purpose
Training Set	318	90%	Used for model training
Validation Set	26	7%	Used for hyperparameter tuning during training
Testing Set	9	3%	Used for final performance evaluation

3.2.3 Data Augmentation

To reduce overfitting and improve model robustness, the following data augmentation techniques were applied.

- 1) Grayscale: Converts images to grayscale to reduce the influence of colour information and emphasise object structures.
- 2) Brightness: Adjusts image brightness to simulate various lighting conditions.
- 3) Exposure: Modifies exposure levels to improve robustness against overexposed and underexposed images.
- 4) Blur: Applies blur effects to simulate motion blur or focus-related distortions.
- 5) Noise: Introduces random noise to simulate real-world image variations and reduce overfitting.

3.2.4 Model Development

- 1) **Model Architecture:** The YOLO11 object detection model was selected due to its high inference speed, suitability for real-time applications, and improved accuracy compared with previous YOLO versions. Two model variants were evaluated: YOLO11n (Nano) and YOLO11s (Small).
- 2) **Hyperparameter Configuration:** The following hyperparameters were used during model training:
 - Epochs: 400
 - Batch Size: 16
 - Image Size: 320x240 pixels
 - Early Stopping: Training was terminated when the validation loss failed to decrease for 100 consecutive epochs (patience = 100).
- 3) **Evaluation:** The performance of the model was evaluated using the following statistical metrics:
 - Precision: Measures the accuracy of positive predictions made by the model.
 - Recall: Measures the model's ability to identify all relevant objects.
 - mAP50: Mean Average Precision at an Intersection over Union (IoU) threshold of 0.5.
 - mAP50-95: Mean Average Precision calculated across multiple IoU thresholds ranging from 0.50 to 0.95 with an interval of 0.05.
 - F1-Score: The harmonic mean of Precision and Recall.

Figure 3

Shows the Loss and mAP Values During YOLO11n Model Training

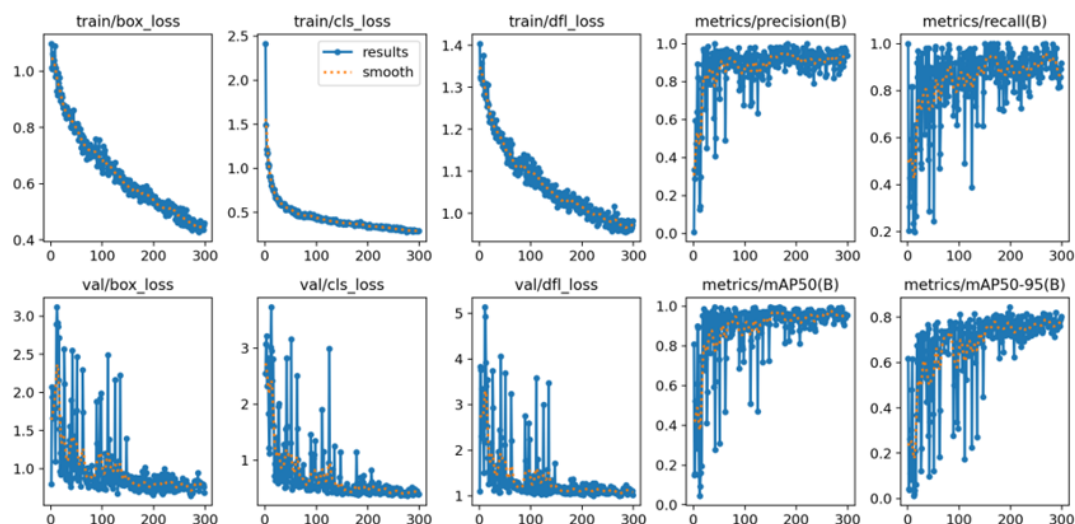
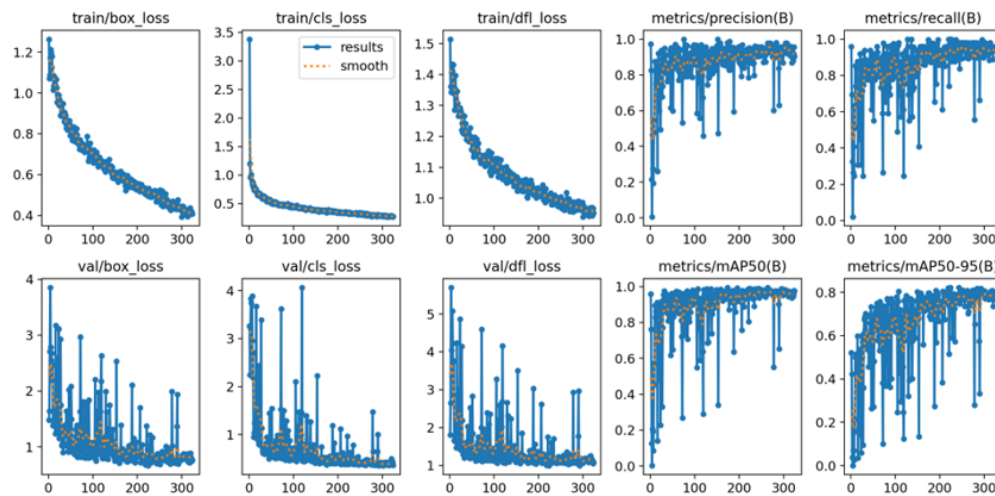


Figure 4

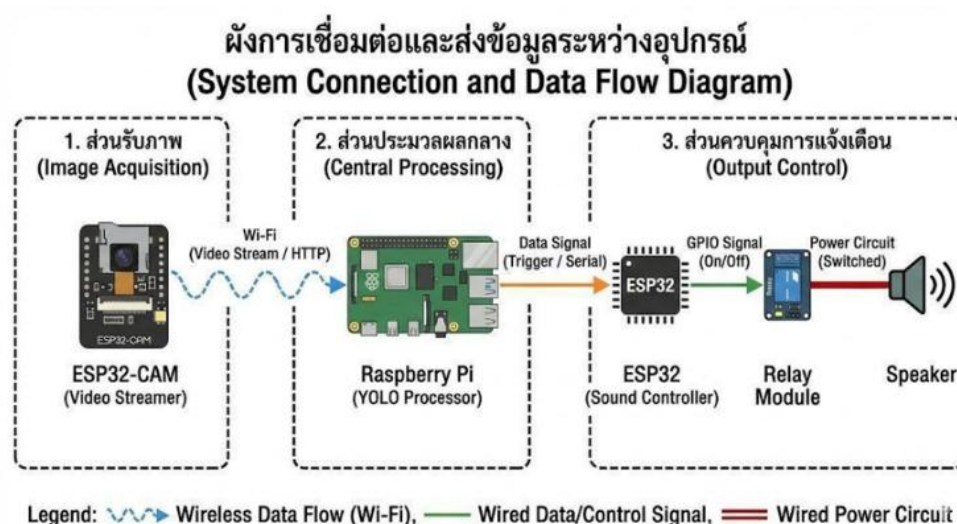
Shows the Loss and mAP Values from the YOLO11s Model Training Data



- 4) **Deployment Implementation:** To enable real-world implementation of the fly detection model within an AI-integrated ultrasonic fly repellent system
- 5) **System Architecture:** The system was designed to operate through the integration of three main components (as illustrated in Figure 2.):
- 6) **Input Node:** An ESP32-CAM module was used to capture and stream video images to a Raspberry Pi through a wireless network.
- 7) **Central Processing Unit:** The Raspberry Pi received image data and performed real-time inference using the YOLO11n model. When a fly was detected, a trigger signal was transmitted to the sound control unit.
- 8) **Output Control Node** A second ESP32 board was used to receive commands from the Raspberry Pi and control a relay module, which activated the speaker system for a predefined duration through a delay-control mechanism.

Figure 5

System Connection and Data Flow Diagram



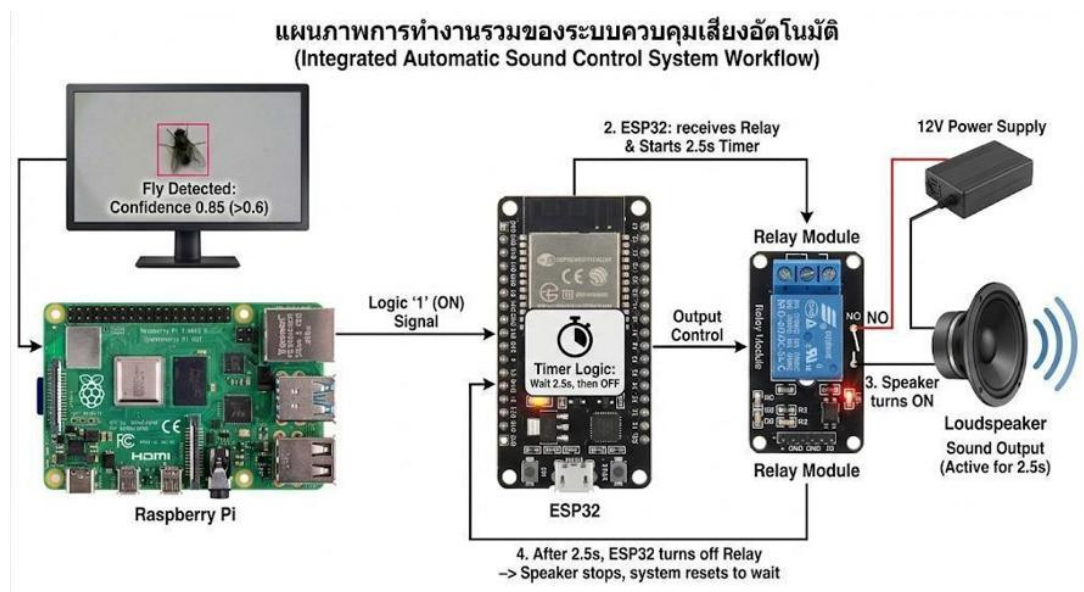
- 9) **Communication Interface:** To ensure coordinated operation among all system components, multiple communication protocols were implemented within the proposed system.
- 10) **ESP32-CAM to Raspberry Pi:** Communication was established through Wi-Fi using the HTTP protocol for real-time video streaming and image transmission.
- 11) **Raspberry Pi to ESP32 (Sound Controller):** Communication was established through Serial Communication (UART) to transmit trigger signals generated by the AI-based fly detection system.
- 12) **ESP32 to Relay Module:** Communication was performed through digital GPIO signals to control the relay switching operation and activate the ultrasonic speaker system.

13) Sound Control Workflow

- a) When the Raspberry Pi detected a house fly (confidence score > 0.5), a trigger signal ("1" or "ON") was transmitted to the ESP32 sound controller.
- b) Upon receiving the signal, the ESP32 activated the relay module (Active High or Active Low configuration).
- c) The relay completed the electrical circuit, enabling the speaker to operate and emit sound.
- d) The ESP32 applied a delay period, allowing the speaker to remain active for 2.5 seconds before automatically deactivating the relay and returning to standby mode for the next detection cycle.

Figure 6

Integrated Automatic Sound Control System Workflow



4. Results

From the experiments and developments above, the following are the results of this study,

4.1 The Experiment Performance of the Ultrasonic Fly Repellent Device

The evaluation of the ultrasonic fly repellent device aimed to identify optimal operational parameters. Experimental results indicated that increasing the ultrasonic frequency from 20 kHz to 35 kHz led to a substantial reduction in house fly landings, decreasing from an average of 13.00 ± 1.73 to 4.33 ± 1.15 individuals, representing a 66.7% improvement in repellent effectiveness. Furthermore, the distance testing demonstrated that the device achieved peak performance when positioned within a 0–5 cm range from the food source (averaging 1.33 ± 1.15 to 2.00 ± 1.00 landings). As the distance increased to 10–15 cm, the effectiveness declined, resulting in an increase in fly landings to 8.67 ± 3.21 and 9.67 ± 3.21 individuals, respectively, due to sound attenuation in the air.

4.2 Performance of the AI-Based Detection System

The integration of the AI-based fly detection system was evaluated based on model accuracy and system reliability. The trained YOLO11n model demonstrated robust detection capabilities, achieving a precision of 0.774, a recall of 0.889, an F1-score of 0.827, and an mAP50 of 0.832, while maintaining an efficient average inference speed of 6.7 milliseconds per image. During integrated system testing, the mechanism exhibited a 95% accuracy rate in speaker activation, with an average response time of 1–2 seconds. These quantitative results confirm that the AI-integrated system provides a reliable, automated, and responsive solution for real-time house fly management.

Beyond the technical development of the AI-integrated ultrasonic fly repellent system, this project has served as a transformative learning experience that bridged academic theory with practical engineering application.

4.3 Connecting Classroom Knowledge to Real-World Innovation

This project provided a unique opportunity to apply knowledge from multiple disciplines, including health sciences, physics, and computer engineering. By integrating my understanding of *Musca domestica* biology and pathogen transmission with principles of sound wave physics such as frequency ranges and sound attenuation I gained a deeper appreciation for how scientific concepts can be translated into public health solutions. This process bridged the gap between the theoretical frameworks studied in the classroom and the complex realities of fly management.

4.4 New Technical Experiences and Computational Skills

Developing the AI-based detection system introduced me to the fields of computer vision and deep learning, which were entirely new areas for me. Mastering the YOLO11 model, from data collection and annotation in Roboflow to training the model and managing the inference workflow on a Raspberry Pi, significantly expanded my computational and problem-solving skills. I learned that building an intelligent system requires not only coding expertise but also a meticulous approach to data preparation and system optimisation.

5. Discussion and Conclusion

The experimental results obtained in this study demonstrate that an AI-integrated ultrasonic fly repellent system is a highly effective, sustainable, and energy-efficient alternative to conventional chemical insecticides. **Analysis of Ultrasonic Fly Repellent Effectiveness** The findings clearly indicate that increasing ultrasonic frequency from 20 kHz to 35 kHz significantly reduces house fly landings. This outcome supports the theory that *Musca domestica* possess specialised sensory organs capable of detecting high-frequency vibrations in air pressure, which directly interfere with their navigation and balance mechanisms. Our observation that 35 kHz yields the highest repellence effectiveness is consistent with previous research, which suggests that frequencies in the range of 30–40 kHz create an unfavourable environment for house flies, inducing natural avoidance behaviours. Furthermore, the significant decline in effectiveness at distances beyond 5 cm underscores the physical phenomenon of sound attenuation in the air. As ultrasonic waves propagate, they lose energy rapidly; therefore, maintaining close proximity to food sources is critical for maintaining a sound pressure level sufficient to trigger an avoidance response in flies. This confirms that the device's placement is a key operational factor for its successful application in real-world food-preparation environments. **AI Integration and Real-Time Detection Efficiency** The development and integration of the YOLO11n object detection model address a significant gap in existing repellent technologies, which often operate continuously and result in unnecessary energy consumption. Our evaluation metrics specifically the mAP50 of 0.832 and an inference speed of 6.7 ms confirm that the system provides high-precision, real-time detection performance even on compact hardware like the Raspberry Pi.

The 95% accuracy in trigger-based speaker activation signifies a robust, automated response system that minimises energy waste by operating only when flies are physically present. By replacing traditional, chemical-based control methods with this intelligent approach, the system aligns with current public health goals to reduce toxic residues in human food chains and environments, thereby mitigating the risk of pesticide resistance in insect populations. Consequently, this system offers a practical, scalable, and environmentally friendly solution for intelligent fly management in both domestic and commercial settings.

In conclusion, the developed AI integrated ultrasonic fly repellent system demonstrated substantial effectiveness based on quantitative experimental results. The system achieved a significant reduction in house fly landings, identified optimal operating conditions in terms of ultrasonic frequency and emitter distance, and successfully integrated an AI-based real-time detection system. Therefore, the proposed innovation shows strong potential for practical implementation as an environmentally friendly alternative to chemical insecticides and supports the development of intelligent fly repellent technologies for household and food preparation environments.

This project has offered a valuable opportunity to apply my classroom knowledge of health sciences and disease vectors to a real-world setting. By developing this system, I have gained a deeper understanding of how interdisciplinary tools, specifically AI can be utilised to address public health challenges. The knowledge I have shared through this project highlights the feasibility of creating low cost, automated solutions to reduce reliance on fly pesticides in households.

Looking forward, this project can be extended to support broader community hygiene. By further refining the hardware and detection algorithms, the system could be adapted for larger, high-traffic areas such as local markets or community food-service centres, where hygiene is essential but often difficult to manage. Ultimately, the most impactful aspect of this work is the potential to shift from reactive fly control to a proactive, environmentally friendly approach. This experience has reinforced my commitment to using technology for social good, and I

hope to continue developing accessible, health-integrated solutions that contribute to cleaner and safer living environments for the wider population.

6. Recommendations

6.1 Policy Recommendations

- 1) Promotion of Chemical Free Food Safety Policies: Given the system's proven 66.7% reduction in fly landings and the critical need to mitigate pesticide resistance, government sectors such as the Ministry of Public Health and local administrative organisations should integrate intelligent ultrasonic repellent systems into standard hygiene certifications, such as the "*Clean Food Good Taste*" programme. This initiative aims to reduce the utilisation of chemical insecticides, minimise toxic residues in the food supply chain, and sustainably elevate public sanitation standards.
- 2) Support for Green Technology and Intelligent Innovation in Academic Research: Educational and research funding agencies should implement national policies that financially support and scale physics and AI-driven (IoT) student innovations. Transforming these academic prototypes into commercial products will foster the development of clean technology and drive the local digital economy.

6.2 Recommendations for Practical Applications

- 1) The prototype ultrasonic fly repellent device can be implemented in households, restaurants, and food-service establishments to reduce house fly activity around food sources and minimise the risk of food contamination.
- 2) Future developments should focus on expanding the operational coverage area, improving the effective working distance of the system, and enhancing product design to support commercial scale production.

6.3 Broader Impact and Knowledge Dissemination

This project serves as a compelling model for other learners and educators on how to address public health challenges through interdisciplinary innovation. By bridging the gap between classroom knowledge in health sciences and practical engineering, this work demonstrates that students can effectively transition from passive learning to active technological intervention. For the wider community, this project highlights the feasibility of creating low-cost, automated solutions that promote sustainable sanitation. I believe that sharing these findings with teachers and peers can inspire further research into health-integrated technology, fostering a culture of problem-solving where students feel empowered to tackle real world environmental and health issues. Ultimately, the most meaningful outcome of this project is the potential to foster a collective commitment to leveraging technology for social good, ensuring that future advancements in science and engineering remain accessible, environmentally responsible, and beneficial to the well-being of the population.

References

- Albert, J. T., & Göpfert, M. C. (2013). Hearing in insects. *Proceedings of the National Academy of Sciences*, 110(38), 15173–15174. <https://doi.org/10.1073/pnas.1307294110>
- Chen, Z., Li, M., & Zhou, W. (2026). Real-time ultrasonic repellent system integration and evaluation. *BioMedical Engineering OnLine*, 25, Article 1517. <https://link.springer.com/article/10.1186/s12938-026-01517-6>
- Department of Disease Control. (2021). *Food and waterborne diseases*. Ministry of Public Health. <https://www.ddc.moph.go.th/dip/news.php?news=17604&deptcode=>
- Eberl, D. F., & Boekhoff-Falk, G. (2002). Development of Johnston's organ in *Drosophila*. *Journal of Neurobiology*, 53(4), 503–511. <https://doi.org/10.1002/neu.10126>
- Heffner, R. S., & Heffner, H. E. (1985). Hearing range of the domestic cat. *Hearing Research*, 19(1), 85–88. [https://doi.org/10.1016/0378-5955\(85\)90100-5](https://doi.org/10.1016/0378-5955(85)90100-5)
- Khamesipour, F., Lankarani, K. B., Honarvar, B., & Kwenti, T. E. (2018). A systematic review of human pathogens carried by the housefly (*Musca domestica* L.). *BMC Public Health*, 18(1), Article 222. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6977770/>
- Liu, W., Yang, M., Yang, G., & Hou, C. (2019). *Musca domestica* as a vector of pathogens. *Journal of Pathogens*, 2019, Article 4851914. <https://doi.org/10.1155/2019/4851914>
- Loh, Y. M. M., & Wilson, R. I. (2015). Acoustic sensing and biomechanical vibration mechanisms in flying insects. *Journal of The Royal Society Interface*, 12(105), Article 20150075. <https://doi.org/10.1098/rsif.2015.0075>
- Raman, V., & Kumar, S. (2017). *Design and development of variable frequency ultrasonic pest repeller* [Unpublished manuscript / Preprint]. ResearchGate. https://www.researchgate.net/publication/318960311_Design_and_Development_of_Variable_Frequency_Ultrasonic_Pest_Repeller
- Sari, N. P., & Wijaya, A. (2020). *Effects of color pattern and sound on the movement of houseflies* [Unpublished manuscript / Preprint]. ResearchGate. https://www.researchgate.net/publication/347304513_Effects_of_Color_Pattern_and_Sound_on_the_Movement_of_Houseflies
- Wag! Editorial Team. (2022, June 3). *Can dogs hear cat deterrents?* Wag Walking. <https://wagwalking.com/sense/can-dogs-hear-cat-deterrents>
- Wang, X., Zhang, Y., & Liu, H. (2022). Physical methods and vector control technologies in food processing environments. *Foods*, 11(14), Article 2035. <https://doi.org/10.3390/foods11142035>



Printed in Thailand

© SEAMEO STEM-ED 2026

SEAMEO STEM-ED

928, 11th Floor, Natural Science Building
Sukhumvit Road, Phra Khanong, Khlong Toei
Bangkok 10110 Thailand

Email: sajse@seameo-stemed.org

Website: <https://seameo-stemed.org/>

