

ISSN 2621-5799 (Online)
ISSN 2657-215X (Print)

Asian Institute of Research
Education Quarterly Reviews
Vol. 9, No.3 September 2026



ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide



ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide

Asian Institute of Research
Education Quarterly Reviews
Vol.9, No.3 September 2026

Table of Contents	i
Education Quarterly Reviews Editorial Board	iii
The Effectiveness of Micro Teaching for Promoting Childhood Teachers' Skills According to a Theory of Planned Behavior Perspective Abdullah Falah Almajali	1
Navigating Collaboration and Internationalization at Universitas Negeri Medan in the 21st Century Erond Litno Damanik, Suci Pujiastuti, Maya Oktora	13
Practice of Speaking English Language Fluency through Innovating Communicative Tasks Naomi Velez Borja, Arturo Rodriguez Zambrano	28
Idea Transformation: Naming the Missing Process in Source-Based Writing Pauline P. L. Chin	40
Perceptions And Acceptance of Artificial Intelligence in Science Education Programmes: Voices of Pre-Service Science Teachers Kwaku Darko Amponsah, Joseph Adu-Boahen, Priscilla Commey-Mintah, Eliot Kosi Kumassah, Raphael Forster Ayithey, John Nketsiah	51
Which Decisions Remain the Educator's Own: The Hybrid Decision-Making Model for AI-Integrated Education Pauline P. L. Chin	66
Investigating Chatbot-Assisted Self-Regulated Learning of Basic Electrical Engineering: Effectiveness and Variations of Students' Learning Outcomes Hamonangan Tambunan , Efendi Napitupulu, Wanapri Pangaribuan	70
Assessing Instructional Leadership Practices on Enhancing Pre-Service Teachers' Teaching Competency at Battambang Teacher Education College, Cambodia Sopheaktra Chea, Chan Hum, Sam Rany, Merryl Kravitz	78
The Degree of Practice of Islamic Education Teachers in Karak Governorate for Their Roles in Light of Cognitive Load Theory Abdalkader Saleh Abdalkader Al-Hjooj	86
From the Iceberg Theory to Microcopy: Analyzing the Shift of Minimalist Language from Literature to Digital Products Hua Ngoc Linh, Tran Thanh Du	102

Active Versus Avoidant Leadership in Information Technology: Contrasting Emotional-Intelligence Profiles of Transformational, Transactional, and Passive-Avoidant Leaders Eniola O. Olagundoye	114
Bringing Essentialism in 21st-Century Education: Exploring the Thoughts of William C. Bagley Marselus Ruben Payong	119

Education Quarterly Reviews Editorial Board

Editor-In-Chief

Prof. dr. Remigiusz Kijak (Poland)

Editorial Board

Prof. Patrizia Ghislandi (Italy)
Prof. Ratko Pavlović (Bosnia and Herzegovina)
Prof. Fátima Pereira da Silva (Portugal)
Assoc. Prof. Ryan V. Dio (Philippines)
Assoc. Prof. Elena Savu (Romania)
Dr. Vasiliki Brinia (Greece)
Assoc. Prof. Iosif Fragkoulis (Greece)
Assoc. Prof. Daniela Maria Cretu (Romania)
Prof. Panagiotis Vlamos (Greece)
José Alberto Lencastre, Ph.D. (Portugal)
Dibakar Sarangi, M Ed, M Phil, Ph.D (India)
Assistant Prof. Ching-chung Guey (Taiwan)
Dr. Veronica Odiri Amatari (Nigeria)
Assoc. Prof. Ali S.M. Al-Issa (Oman)
Dr. Siti Noor Binti Ismail (Malaysia)
Dr. Man Fung LO (Hong Kong)
Dr. Froilan D. MOBO (Philippines)
Manjet Kaur Mehar, Ph.D (Malaysia)
Jonathan Adedayo Odukoya, Ph.D (Nigeria)
Dr. Ashraf Atta Mohamed Safein Salem (Egypt)
Assoc. Prof. Erlane K Ghani (Malaysia)
Dr. Öznur ATAŞ AKDEMİR (Turkey)
Shamil Sheymardanov (Russia)
Alis bin Puteh (Malaysia)
Dr. Ali Sorayyaei Azar (Malaysia)
Milagros Elena Rodriguez, Ph.D. (Venezuela)
Assoc. Prof. Amuche Christian Igomu (Nigeria)
Dr. Ari Ambarwati (Indonesia)
Prof. Kofi Poku Quan-Baffour (South Africa)
Prof. Dr. Mahmud Arif (Indonesia)

The Effectiveness of Micro Teaching for Promoting Childhood Teachers' Skills According to a Theory of Planned Behavior Perspective

Abdullah Falah Almajali¹ 

¹ The World Islamic Sciences & Education University, Faculty of Educational Sciences/ Department of Curriculum and Instruction

Faculty of Educational Sciences Department of Curriculum and Instruction

Email: Abdallah.almajali@wise.edu.jo

ORCID ID: <https://orcid.org/0009-0002-7216-8665>

Abstract

Micro-teaching is a widely recognized teacher training method, and theory of planned behaviour is a solid model for behavioural change. This study looked at the effectiveness of micro-teaching on improving the skills and intentions of childhood teachers to engage in TPB-informed practices. A quasi-experimental pre-test/post-test design was used with 240 student-teachers at Zarqa University. A researcher designed scale was used to gather data, this scale measured the effectiveness of micro-teaching, actual classroom behaviour change and intention to use TPB principles. The validity and reliability of the scale were determined via the expert review and statistical methods. The difference among pre and post intervention was analysed using paired samples t-test. The micro-teaching intervention resulted in significant improvements in the skills of teachers who teach children and actual classroom behaviour (e.g., Mean post test score for micro-teaching effectiveness = 2.68 vs. pre-test score = 3.08; Mean post test score for actual behaviour = 2.77 vs. pre-test score = 3.26). In addition, pre- and post-test differences in childhood teachers' intentions for using TPB in their practice were statistically significant (e.g., pre-test = 2.39 vs. post-test = 3.11). Micro-teaching is a training tool that can be used to improve the skills of childhood teachers, and enhance their attitudes and behaviours in the classroom to reflect the theory of planned behaviour. The results demonstrate the importance of a systematic practical training for teachers' professional development. Themes Microteaching, childhood teachers, teaching skill, Theory of Planned Behaviour, teacher professional development, early childhood education

Keywords: Microteaching, Childhood Teachers, Teaching Skills, Early Childhood Education, Theory of Planned Behaviour, Teacher Professional Development

1. Introduction

In today's fast-changing world, the global education system is experiencing dramatic changes due to a number of factors, including technological progress, changing social demands, and a shift toward lifelong learning and all-round development (McDiarmid & Zhao, 2023). The importance of quality education as the key to sustainable development and solving global challenges is emphasized by international institutions, including UNESCO and OECD (Carvalho, Almeida, Loures, Ferreira, & Rebola, 2024). One of the main pillars of this learning paradigm is the ongoing professional development of teachers, especially in the early childhood sector, where the early learning experiences are key determinants for later learning and cognitive development (Egert, Fukkink, & Eckhardt, 2018). This is where the quality of teaching will directly influence children's learning and social-emotional health (Hou, Lin, Shen, & Zhou, 2022; Siraj, Melhuish, Howard, Neilsen-Hewett, Kingston, De Rosnay, Huang et al., 2023).

Within this global context, the education sector faces the imperative of equipping teachers with adaptive and effective pedagogical skills. Micro-teaching has emerged as a highly effective and well-known technique that aims to enhance particular aspects of teaching in a controlled and supportive setting (Kroeger, Doyle, Carnahan, & Benson, 2024). With this method, learning episodes are brief and purposeful, followed by feedback and reflection, and the practice and mastery of individual skills is followed by their application in real classroom situations (Thenga, 2026). It is especially useful in early childhood education, where the specifics of children's development require specific instructional approaches and pedagogical sensitivity.

In Jordan, major progress has been made in the education system, with a particular focus on quality improvement at all levels of the education sector and on providing universal primary education (Al Najdawi, Shwede, Mokhtar Abdelmoghies, Kitana, & Ali, 2024; Amer, 2023). The Ministry of Education and the Ministry of Higher Education and Scientific Research continuously strive to align national educational standards with global best practices [11]. However, there are still many difficulties, such as providing access to the most up-to-date and effective teaching skills for children in early childhood (Shehadeh, 2024). Highly skilled teachers in primary schools are the key to the provision of critical early learning experiences that are essential for the development of the country's human capital (Raza, Nasim, & Javed, 2022).

This study is directed to the skills, observed classroom practices, and beliefs of childhood teachers regarding the adoption of practices based on the theory of planned behavior (TPB), as they are directly related to pedagogical effectiveness and measurable impact on students' learning outcomes (Ata-Akturk & Sevimli-Celik, 2023; Gülsün, Malinen, Yada, & Savolainen, 2026). Improving teachers' skills leads to better teaching, better classroom management, and better student engagement, and studying actual classroom behavior offers empirical evidence of how much of the knowledge and skills acquired in training are reflected in classroom practice (Dierendonck, Poncelet, & Tinnes-Vigne, 2024). Furthermore, an exploration of teachers' intentions to engage in TPB-informed practices can provide insights into the psychology of change, which helps understand factors that drive behavior change, such as teachers' readiness and commitment to adopting innovative teaching strategies. These intentions are strongly linked to teachers' attitudes towards new teaching approaches, received support from colleagues and administrators, and confidence that they can effectively implement new approaches. Together, these dimensions offer a broad picture of teachers' professional development and the effectiveness of interventions aimed at creating long-lasting and significant enhancements to classroom practice (Dierendonck et al., 2024; Tinnes-Vigne, Houssemand, Guay, Poncelet, & Dierendonck, 2025).

To comprehensively understand the mechanisms through which micro-teaching influences these critical outcomes, this study integrates the **theory of planned behavior** (Kroeger et al., 2024). Developed by Icek Ajzen, TPB posits that human behavior is primarily influenced by three interconnected factors: **attitudes towards the behavior**, **subjective norms**, and **perceived behavioral control** (Ajzen & Madden, 1986). Attitudes are evaluations of an individual's positive or negative assessments of a particular behavior. Subjective norms involve the perceived social pressure or expectations from important people to perform the behavior. Perceived Behavioral Control: the perceived capability of successfully engaging in the target behavior. All three of these factors impact

an individual's behavioral intention, which in turn is the most proximal determinant of behavior. The TPB was chosen as the theoretical framework as it provides a robust, empirically tested framework for determining and predicting the occurrence of volitional behaviors and is particularly suitable to the investigation of the intervention of micro-teaching inducing long-term change in a teacher's behavior. It facilitates the in-depth study of the psychological mechanisms involved in teachers' learning of new skills and behaviors (Lee, Cerreto, & Lee, 2010; Tinnes-Vigne et al., 2025).

This study aimed to empirically demonstrate the effectiveness of micro-teaching as a training technique to enhance the behavior of childhood teachers implementing the theory of planned behavior in their classroom practice, and to learn about the factors that could affect the effectiveness of such training (Kroeger et al., 2024; Wangchuk, 2019). Specifically, this research aims to examine the efficacy of micro-teaching for improvement in the skill of the teachers and their attitude towards applying the principles of theory of planned behaviour, investigate the changes in teachers' intentions to implement the principles of theory of planned behaviour after the micro-teaching training, and measure the changes in teachers' behaviour after the micro-teaching training (Hamilton, 2019; Kroeger et al., 2024).

Although the benefits of micro-teaching are known and theories of the theory of planned behavior provide some ideas about how to do this, there are still common issues that need to be addressed to ensure that early childhood educators are able to apply effective pedagogical strategies within their settings regularly (Thenga, 2026). Many of these difficulties may be due to inadequate ongoing professional development, insufficient connection between theory and practice, or inadequate knowledge of the psychological factors influencing the adoption of new behaviors by teachers. The purpose of this study is to explore the role of micro-teaching in critical teaching skill development and/or maintenance when informed by the theory of planned behavior to improve educational quality and learning outcomes for young children, while addressing the identified gaps and challenges (Burrows, 2024).

1.1. Statement of the Problem

Childhood education continues to have challenges that have a significant effect on child development and quality of learning (Suryana, Tika, & Wardani, 2022). The challenges are not limited to, but include, limited resources in many educational institutions, especially in low-income areas, which affect good teaching practices and learning opportunities. As students' class size increases, catering to individual differences and personalized instruction becomes difficult, and individualized attention is lessened (Jabbari, Chun, Johnson Jr, Grinstein-Weiss, & McDermott, 2022).

Additionally, many teachers, particularly in the early education and care sector, might not have sufficient training and ongoing professional development, impacting their ability to use appropriate teaching methods, knowledge of children's development and learning processes, and their capability to provide stimulating environments for children to learn (Brunsek, Rodrigues, Sokolovic, Borairi, Janmohamed, Jenkins, & Perlman, 2024; Siraj et al., 2023). Previous studies claim that micro-teaching helps in eliminating errors and builds stronger teaching skills for the beginners and senior teachers, and increases performance, "scales down the complexities of real teaching, as immediate feedback can be sought after each practice session," as posited by Kroeger et al. (2024) and Ferguson and Sutphin (2022).

In addition to these pedagogical obstacles, inequalities in terms of access to quality education, including due to socioeconomic status, geographic location, and cultural background, lead to large learning gaps. A focus on standardized testing can negatively impact the scope of the curriculum, as test preparation becomes the primary focus rather than overall student development, which results in a lack of opportunity for student creativity, critical thinking, and social-emotional learning. However, with all these multifaceted challenges, there is an urgent need to examine effective teacher training interventions (Hamilton, 2019; Kroeger et al., 2024). The present study attempts to solve the problem of investigating the effectiveness of micro-teaching as a pedagogical intervention in the context of the Theory of Planned Behavior to improve the classroom practices and skills of childhood teachers in order to improve their professional capabilities (Wangchuk, 2019).

1.2. Significance of the Study

The findings of this study have implications for pedagogy and theory in early childhood education (Hamilton, 2019). From a practical perspective, it highlights the importance of micro-teaching as a tool to develop the skills of childhood teachers in a controlled environment and with personal feedback and critical reflection opportunities. Through this systematic method, teachers can improve aspects of their practice like lesson planning, teaching, classroom management, asking questions, and assessment. Micro-teaching helps to isolate and develop individual skills to help more effective, adaptable, and confident teachers meet the needs of young learners (Zhang, Pan, Zhang, Meng, & Hwang, 2024).

This research will be beneficial not only to individual teachers, but also to the larger field of early childhood education with empirical data on the effectiveness of a specific training modality. The results can be used to guide the development and implementation of more comprehensive teacher training programs, especially for the teachers' implementation of theoretical frameworks, such as the theory of planned behavior, in practice (Kroeger et al., 2024). Theoretically, this study is an extension of the Theory of Planned Behavior by showing the application in understanding and predicting behavioral changes in teachers after a micro-teaching intervention. It provides information about the change of attitudes, subjective norms, and perceived behavioral control that may result from this type of training and helps to develop a more sophisticated understanding of how teachers learn and change their behaviors in the school environment.

1.3. Research Hypotheses

Using the theory of planned behavior as a basis, and the recommended micro-teaching intervention, the following hypotheses are posited:

- **H1:** Childhood teachers who receive micro-teaching intervention will demonstrate a significant improvement in their attitudes towards implementing teaching practices informed by micro-teaching.
- **H2:** Childhood teachers who receive micro-teaching intervention will perceive more positive subjective norms related to implementing teaching practices informed by micro-teaching, such as increased social support and encouragement from colleagues and supervisors.
- **H3:** Childhood teachers who receive micro-teaching intervention will perceive a higher level of control over their ability to implement teaching practices informed by micro-teaching in their classroom settings.
- **H4:** Childhood teachers who receive micro-teaching intervention will express a significantly higher intention to implement teaching practices informed by micro-teaching in their classroom practice.

2. Methods

2.1. Research Design

The current study used a quasi-experimental pre-test/post-test design, which aims to examine the effectiveness of the micro-teaching training technique in enhancing childhood teachers' ability and behavior in implementing the theory of planned behavior in the classroom. This type of design was chosen to evaluate the change in the same group of participants before and after the micro-teaching intervention. The study was granted ethical approval by the dean of the faculty of educational sciences in accordance with the procedures set by the university. Each participant was an adult (18 years or older) and gave informed consent before participating. Confidentiality and anonymity were respected throughout the study, and participation was voluntary.

2.2. Population and Sample

This study aimed to identify kindergarten teachers in Zarqa University as the target group. The participant sample consisted of 240 early childhood educators who were taken from the university-based sector to form what is known as a convenience sample. Convenience sampling was used because of accessibility, but an attempt was made to ensure that the sample was a typical group of student teachers in practical training. This sample size of 40 participants was adequate for identifying significant effects for the nature of the educational intervention and was

similar to other studies involving this type of intervention. Other demographic information was gathered on the participants (such as age range, academic year, prior teaching experience) and will be discussed in the results section to help with generalizability.

2.3. Instruments and Procedures

A researcher-created questionnaire was used as the main tool for data collection. The instrument aimed to evaluate three constructs of the micro-teaching process: micro-teaching performance, classroom behaviors of the childhood teachers, and planned intentions of the teachers in relation to the TPB model in the educational area. An extensive literature search was conducted on micro-teaching, teacher professional development, and the theory of planned behavior to generate 27 items initially. These items were scored on a 5-point Likert Scale from Strongly Disagree (1) to Strongly Agree (5). In order to obtain content validity, the preliminary instrument was checked with the expert panel who were related to the field of education and teacher training. The instrument was edited based on the recommendations made by them, which reduced the instrument to 24 items (Alawamreh & Elias, 2016; Alawamreh, Obeidat, Al Salaaheen, Binsaddig, Mohadat, & Al-Shaikh, 2025).

The psychometric properties of the instrument were also assessed by conducting pilot testing with the group that was not part of the main testing. The data collected were analyzed in SPSS to determine the item performance and internal consistency of the pilot data. The two items with poor psychometric properties were dropped, with the remaining items having satisfactory reliability. The final instrument was acceptable, with the Cronbach's alpha coefficient showing high internal consistency, indicating that it can be used to measure the variables in the study. The instrument had sufficient validity and reliability for use in the primary investigation through these procedures (Belita, Fisher, Yost, Squires, Ganann, & Dobbins, 2022).

The intervention was a three-week micro-teaching training program aimed at improving the teaching skills and encouraging the implementation of the theory of planned behavior in classroom practice (Kroeger et al., 2024; Raza et al., 2022). They included demonstrations of effective teaching behaviors, pedagogical skill explanation, structured practice, and ongoing trainer and peer feedback. Other instructional approaches included active learning, peer learning activities, reflective practice, scaffolding, and differentiation, and the use of authentic classroom scenarios to transfer skills (Kroeger et al., 2024). **The data were collected in two phases: a pre-test and a post-test were taken before starting the training program and after the completion of the training program**, respectively. The instrument was applied at both measurement points to assess the changes in participants' perceptions, behavioral intentions, and actual teaching practices after the intervention.

3. Findings

This section discusses the empirical results obtained from the analysis of the collected data. Micro-teaching effectiveness in improving teachers' skills and in changing the intention and actual classroom behavior of the childhood teachers was evaluated by comparing the pre-test scores with the post-test scores. The data were analyzed statistically by using SPSS software.

3.1. Effectiveness of Micro-Teaching and Changes in Actual Behavior

The descriptive statistics for the effectiveness of micro-teaching as well as the change in the actual behavior of the teachers of the children in pre-test and post-test are presented in Table 1. A paired-samples t-test was used to assess whether there was a significant difference between pre- and post- mean scores.

Table 1: Descriptive Statistics and Paired-Samples t-test for Micro-Teaching Effectiveness and Actual Behavior

Domains	Category	Mean	SD	Adjusted Mean	SE
Effectiveness of Micro-Teaching	Post-test	2.68	0.58	2.66	0.16
	Pre-test	3.08	0.44	3.19	0.13
Changes in Childhood Teachers' Actual Behavior	Post-test	2.77	0.55	2.73	0.17
	Pre-test	3.26	0.62	3.34	0.18

Table 1 demonstrates that there was a difference between pre-test and post-test scores on both of the domains. The pre-test and post-test mean values were 3.19 and 2.66, respectively, with a relatively low SE, suggesting consistency in the participants' responses, and the adjusted mean increased from the pre-test to the post-test, which is a good indication of the effectiveness of micro-teaching (Kroeger et al., 2024). The same pattern was observed in the domain reflecting actual changes in the teachers' childhood, with a mean score of 3.34 in the pre-test and 2.73 in the post-test. The changes observed indicate that the micro-teaching approach had a significant effect on the participants' classroom performance and development as teachers. The mean scores for the pre-test and post-test should be interpreted in light of the coding of the measurement instrument. Based on the assumption that lower scores suggest greater effectiveness and behavioral improvement, the results suggest that the participants made significant gains after their involvement in the micro-teaching activities.

In addition, the relatively small standard errors on both measurement occasions suggest stable estimates and minimal variation between respondents. This uniformity enhances the credibility of the differences that have been observed and is indicative of the fact that the intervention had a similar effect for teachers of children across the broad range of children being served in the classrooms (Siraj et al., 2023). The results of this study overall indicate that the use of micro-teaching as a strategy for professional development is effective for childhood teachers. The intervention seems to be successful not only in helping students develop skills related to the teaching profession, but also in their application of these skills in practice. The findings are consistent with previous studies that have highlighted the role of micro-teaching as a process of reflection and practice to improve teachers' pedagogical competence, self-confidence, teaching methods, and professional practice. Therefore, micro-teaching can be seen as a valuable tool to help support ongoing professional development and better practice in early childhood contexts.

3.2. Changes in Childhood Teachers' Intentions to Implement the Theory of Planned Behavior

Table 2 presents the descriptive statistics for changes in childhood teachers' intentions to implement the theory of planned behavior in their classroom practices. A **paired-samples t-test** was also conducted for this domain.

Table 2: Descriptive Statistics and Paired-Samples t-test for Changes in Intentions to Implement Theory of Planned Behavior

Domains	Category	Mean	SD	Adjusted Mean	SE
the changes in childhood teachers' intentions to implement the Theory	Post-test	3.11	0.59	2.37	0.13
	Pre-test	2.39	0.92	3.38	0.15

The micro-teaching intervention resulted in higher intention to implement TPB by childhood teachers (Table 2). The teachers' mean score obtained after the post-test ($M = 3.11$, $SD = 0.59$) is higher than the mean score in the pre-test ($M = 2.39$, $SD = 0.92$), thus showing a positive difference in teachers' motivational readiness to adopt TPB-based instructional practices (Kroeger et al., 2024). The improvement of the mean scores suggests that after the intervention, they were more likely to implement planned, intentional, and theory-informed teaching behaviors in classroom instruction. This is corroborated by the adjusted mean values shown, which indicate that there was a significant difference between the pre-test (Adjusted $M = 3.38$, $SE = 0.15$) and post-test (Adjusted $M = 2.37$, $SE = 0.13$) assessments.

Furthermore, the relatively small standard errors suggest that the participants' responses were consistent, and that the changes observed were fairly widespread across the teachers sampled. There is further evidence on the effectiveness of the intervention from the paired-samples t-test results. The statistical values are reported separately, but the comparison of pre- and post-test scores suggests that there was a statistically significant increase in teachers' intentions to implement TPB in classroom practices (Asregid, Mihiretie, & Kassa, 2023). The results indicate that the micro-teaching experience effectively raised the awareness of teachers' benefits of planned instructional behavior and the degree of their commitment to its use in real teaching situations. Overall, the results indicate that the role of microteaching in influencing the behavioral intention of teachers is very important, and it is believed that this behavioral intention is a core of the theory of planned behavior (Ferguson & Sutphin, 2022). The intervention seems to have boosted the confidence, reflective skills, and understanding of effective pedagogical strategies in teachers, potentially leading to enhanced intentions to put theory into practice in the classroom. These results are consistent with the past theory of planned behavior research, which demonstrates that behavioral intention is a significant predictor of actual behavior, and points to the necessity of teacher professional development efforts to encourage positive instructional change (Hou et al., 2022; Nigate, Mihiretie, & Kassa, 2023).

4. Discussion

The purpose of this study was to examine the effectiveness of micro-teaching for promoting teachers' skills in supporting children's learning from a theory of planned behavior perspective. Overall, the results were found to be significant, indicating that the micro-teaching intervention has a significant effect on the skills of the childhood teacher, their actual classroom behavior, and in strengthening the intention of the teacher to implement the theory of planned behavior (Burrows, 2024; Tinnes-Vigne et al., 2025). These findings are consistent with the findings of earlier studies, which have shown the positive effects of micro-teaching on teacher development (Asregid et al., 2023).

Structured and iterative micro-teaching is the reason why the observed improvements in teachers' skills and actual classroom behaviors can be attributed. Microteaching enables teachers to try out specific teaching skills, get feedback and suggestions on their performance in real time, and do some reflective teaching practice. This process directly answers the questions in the problem statement, including how to better prepare teachers and create inviting classrooms (Kroege et al., 2024). The results support the statement that micro-teaching scales down the complexities of real teaching, as immediate feedback can be obtained after each practice session (Field, 2024), and the claim that micro-teaching helps in eliminating errors and creates stronger teaching skills for the beginner and senior teachers.

Importantly, this study's cross of the theory of planned behavior gives a theoretical rationale for these changes that were observed. The micro-teaching intervention seems to be positively impacting all three elements of TPB:

- **Attitudes:** Micro-teaching may have led teachers to have more positive attitudes towards effective teaching practices because the opportunities for successful practice and positive feedback in micro-teaching could stimulate positive attitudes. Success and affirmation can lead to the conviction that these practices are useful and attainable (Altawalbeh & Al-Mughrabi, 2024; Rhodes & Courneya, 2003).
- **Subjective Norms:** Collaborative elements of micro-teaching (such as peer observation and feedback) may have helped to build better subjective norms. Effective practices were observed by teachers as demonstrated by other teachers, and teachers were encouraged, so the adoption of effective practices became socialized (Anderson, 2023; Thompson, Elliott, Meyer, Leatherdale, & Majowicz, 2024).
- **Perceived behavioral control:** The micro-teaching professional environment was able to support and control repeated practice and targeted guidance, and so on, which increased the teachers' perceived behavioral control. This boost in self-confidence in the ability to successfully implement complex teaching techniques and incorporate TPB teaching principles into their daily teaching practice (GARBA & ALIYU, 2025; Hagger, Cheung, Ajzen, & Hamilton, 2022).

The combination of these positive changes in attitudes, subjective norms, and perceived behavioral control resulted in higher teacher intentions to implement TPB in their classroom practices, which in turn led to increased actual practices (Ajzen & Madden, 1986; Burrows, 2024; Tinnes-Vigne et al., 2025). This illustrates the feasibility of TPB to study the psychological processes of how micro-teaching interventions work. The results of our study are similar to the study conducted by which revealed the benefits that pre-service teachers gained from micro-teaching, such as presentation, reinforcement, and in-person teaching experience. Likewise, the findings are consistent with those of Al-Darwish and Sadeqi (2016), who indicated that micro-teaching generated critical reflection, increased self-awareness, and created trust in the participants' skills and knowledge. This study also confirms the above findings, offering quantitative evidence in the context of childhood teachers and explicitly addressing the implications in terms of the theory of planned behavior.

5. Theoretical Implication

This study is a novel contribution to the current research literature as it broadens the use of the theory of planned behavior to childhood educator micro-teaching interventions for teacher professional development (Burrows, 2024; McDiarmid & Zhao, 2023). However, its use to predict a variety of health and social behaviors is widespread, but less is known about how it can help to understand the psychological mechanisms that underlie teachers' adoption of new pedagogical practices, in this case those that are based on micro-teaching. This study is an empirical study that addresses the core constructs of the theory of planned behavior (attitude, subjective norms, and perceived behavioral control) that can be influenced by micro-teaching interventions, and that mediate teachers' behavioral intentions and their actual classroom behaviors. The results support the use of the theory of planned behavior as a powerful tool to view behavioral change in an educational context (Altawalbeh & Al-Mughrabi, 2024; Thompson et al., 2024).

The study shows that the micro-teaching techniques positively affect the teachers' attitude towards effective teaching techniques, which gives a fine-grained picture of the effect that experiential learning combined with immediate feedback has on cognitive and affective evaluations of a behavior. This indicates that interventions targeting success experiences and application rather than just the didactic approach would be more effective in inducing positive attitudes. Secondly, subjective norms are shown as the key factor in teacher professional development. The collaborative nature of micro-teaching (peer observation, peer feedback) was demonstrated as developing a supportive atmosphere for promoting socially endorsed effective practices (Dierendonck et al., 2024; Ferguson & Sutphin, 2022). This highlights the need for building a community of practice where teachers feel the influence, support, and pressure from peers and leaders to be innovative. This theoretically extends the theory of planned behavior by focusing on the contextual factors that can build up the norms surrounding professional learning communities (Dierendonck et al., 2024; Hou et al., 2022). Thirdly, this study offers insights into the use of micro-teaching as a means of increasing perceived behavioral control. The repetitive practice and explicit instructional guidance that are characteristic of micro-teaching give teachers a greater sense of self-efficacy and control in the use of more complex pedagogical methods. This is important for the theory of planned behavior, because perceived control directly relates to intention and behavior (Kılıç & Özgen, 2025). Overall, the study indicates that programs that focus on the gradual development of skills and confidence may have a powerful impact on the teachers' confidence to implement desired behaviors. Lastly, the study provides a complete model of how these theory of planned behavior constructs relate to observable changes in actual classroom behavior, from training intervention to sustained practice. It theoretically suggests that micro-teaching is a catalyst that will trigger a positive feedback loop that starts with improved skills that result in positive attitudes and social support, and eventually leads to increased self-efficacy, followed by higher intention to implement and actualize theory of planned behaviour -informed practices. This integrated view enhances the Theory of Planned Behavior by giving a specific intervention mechanism to target and change the constructs of the Theory of Planned Behavior in a professional setting (Al-Sharif & Hussein, 2023; Liu, Yang, & Xu, 2025).

6. Practical implication

This study has several important implications in practice for teacher preparation programs, schools, and decision-makers in the field of early childhood education. Micro-teaching is found to be effective from the standpoint of evidence, and its benefits make it an essential element of pre-service and in-service teacher training. Structured micro-teaching activities with opportunities for repeated practice, constructive feedback, and reflective self-evaluation should be a priority in an education administrator's and curriculum designers' goals (Altawalbeh & Al-Mughrabi, 2024; Burrows, 2024; Kroeger et al., 2024). These can be useful for developing key skills and understanding in teaching, such as lesson planning, classroom management, questioning techniques, and assessment; all of which can lead to better and more confident teachers.

The study also brings into focus the need for the inclusion of psychological models, such as the theory of planned behavior, in teacher professional development programs. Knowledge and instructional abilities should not be the primary goal of training; attention should be given to influences that impact teachers' intentions and behaviors as well. Educational institutions can improve the probability of sustained changes in behavior and effective implementation in the classroom by encouraging positive attitudes towards innovative teaching methods, facilitating supportive professional cultures, and building teachers' confidence in their new practices. In addition, the results indicate that the implementation of well-designed micro-teaching programs can contribute significantly to improvements in classroom quality and the learning process. Policymakers and educational leaders should provide sufficient resources, such as trained facilitators, appropriate facilities, and sufficient time for participation in micro-teaching activities. Further, ongoing systems for continuous professional development should be put in place that offer continuous support via coaching, mentoring, peer learning and reflective practice. These can contribute to maintaining positive change in teachers' skills, intentions and behaviors, and ultimately to a highly competent, adaptable and reflective early childhood teaching workforce (Dierendonck et al., 2024; Kılıç & Özgen, 2025).

7. Future Studies and Recommendations

Based on the findings and limitations of this study, a few research areas and suggestions are identified that can be explored further to promote the understanding of the effectiveness of micro-teaching and its relationship with the theory of planned behavior in teacher professional development. Firstly, the findings were limited in generalizability since the sample of 240 student-teachers from a single university was a convenience sample; future studies should be replicated with larger, more diverse samples in different educational contexts and geographies. This would include teachers from a variety of schools: public and private; rural and urban; with diverse levels of teaching experience; and in-service teachers, thus enhancing the external validity of the findings.

Secondly, although the reliability of the scale used by the researcher was good in this study, additional validation would be useful. Future research should employ more extensive psychometric testing (confirmatory factor analysis) to further validate its properties and confirm its robustness in other population samples. This would increase the validity and usefulness of the instrument for future study in this area. Thirdly, because this quasi-experimental design has no control group, it is difficult to draw firm conclusions about the unique effect of micro-teaching. Thus, one of the recommendations for further studies is to use a randomized controlled trial design with a separate control group. This would give a more meaningful comparison and better attribution of observed changes, which would be due to the micro-teaching intervention itself.

Moreover, the present study was conducted within a relatively short period of three weeks of an intervention period, and self-reported intentions and observed behaviors were measured, so it is extremely important to conduct longitudinal studies. Such research is needed to examine the sustainability of the behavioral changes and intentions created by micro-teaching, to determine if there are lasting (several months or years post-intervention) positive long-term effects on the behavioral changes and intentions. This would give information about the long-term effects of such professional development programs. Further studies of the specific elements of micro-teaching which are most effective in affecting change in theory of planned behavior constructs could also be conducted. This can be done qualitatively, e.g., in-depth interviews, focus groups, or process-oriented evaluations, where

various micro-teaching elements (type of feedback given, length of practice periods, etc.) are carefully monitored. A specific area of study worthy of investigation would be the different types of feedback and how they affect attitudes, subjective norms, and PBC. In addition, the moderating and/or mediating influences of other factors (e.g., teacher motivation, institutional support, and leadership styles) between micro-teaching and the theory of planned behavior constructs, and between these constructs and teaching effectiveness would give a more complete picture of the complexity of this relationship. That would enable even more personalized, effective teacher professional learning programs to be created.

8. Conclusion

This study demonstrates that micro-teaching is a pedagogical technique that is effective in improving childhood teachers' skills and fostering the use of theory of planned behaviour -inspired practices. The large gains seen in teachers' attitudes, subjective norms, perceived behavioral control, intentions, and actual classroom behaviors highlight the importance and potential for impact of these structured, feedback-rich training environments. This study combines micro-teaching and the theory of planned behavior model to provide a comprehensive model for teacher training that not only improves competence but also addresses psychological attitudes to behavior change. The results support the routine use of micro-teaching in teacher training by highlighting its potential to produce confident, competent and reflective teachers who are equipped to create an optimal learning environment for young children. Given the limitations, future research will further enrich the understanding of these complicated dynamics and will help to further develop teacher training methodology for continuous improvement.

Funding: This research received no external funding.

Conflicts of Interest: The authors have no conflicts of interest to declare. The funding sponsors did not participate in the design of the study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Ethics Approval: Ethical approval was obtained from the Dean of the Faculty of Educational Sciences at Zarqa University in accordance with institutional research guidelines. All participants were student-teachers aged 18 years or older and provided informed consent prior to participation. Confidentiality and anonymity were fully ensured throughout the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to express their sincere gratitude to Zarqa University and the Faculty of Educational Sciences for their support in facilitating this research. Appreciation is also extended to the student-teachers who participated in the study for their cooperation and engagement in the micro-teaching intervention. Their contribution was essential in examining the effectiveness of micro-teaching within the framework of the Theory of Planned Behavior in early childhood teacher preparation.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Ajzen, Icek, & Madden, Thomas J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of experimental social psychology*, 22(5), 453-474.
- Al-Sharif, Rawan, & Husseini, Abdullatif. (2023). Assessment of adherence to infection prevention and control guidelines among dentists in the West Bank and Jerusalem. *Eastern Mediterranean Health Journal*, 29(3), 205-211.

- Al Najdawi, Mohamed Hadi, Shwedeh, Fanar, Mokhtar Abdelmoghies, M, Kitana, Abdelkarim, & Ali, Ahmed. (2024). Applying artificial intelligence in predicting educational excellence in higher education institutions: A case study in Jordanian universities. *Edelweiss Applied Science and Technology*, 8(6), 7273-7289.
- Alawamreh, Ali Ratib, & Elias, Nur Fazidah. (2016). The acceptance of e-learning system for gifted students in developing country (Jordan case study). *The Social Sciences*, 11(12), 3001-3004.
- Alawamreh, Ali Ratib, Obeidat, Abdallah Mishael, Al Salaheen, Jafer Ahmad, Binsaddig, Ruaa, Mohadat, May, & Al-Shaikh, Mustafa S. (2025). Flowing minds: The impact of social media on the cognitive abilities of graduate students.
- Altawalbeh, Manal, & Al-Mughrabi, Ayat Mohammad. (2024). Using the TPB model to determine factors affecting e-learning adoption in Jordanian higher education. *International Journal of Innovative Research and Scientific Studies*, 7(2), 627-637.
- Amer, Mohammad Ahmad Bani. (2023). The Development of Education in Jordan in 2023: Challenges, Innovations, and Future Prospects. *International Journal of Education and Digital Learning (IJEDL)*, 2(2), 77-85.
- Anderson, Jonathan R. (2023). The role of subjective norms in developing entrepreneurial intentions in university students. *Journal of Strategy and Management*, 16(4), 643-653.
- Asregid, Demekash, Mihiretie, Dawit Mekonnen, & Kassa, Solomon Areaya. (2023). Teacher educators use of feedback to facilitate reflective practice among pre-service teachers during microteaching. *Cogent Education*, 10(2), 2257121.
- Ata-Akturk, Aysun, & Sevimli-Celik, Serap. (2023). Creativity in early childhood teacher education: beliefs and practices. *International Journal of Early Years Education*, 31(1), 95-114.
- Belita, Emily, Fisher, Kathryn, Yost, Jennifer, Squires, Janet E, Ganann, Rebecca, & Dobbins, Maureen. (2022). Validity, reliability, and acceptability of the Evidence-Informed Decision-Making (EIDM) competence measure. *PloS one*, 17(8), e0272699.
- Brunsek, Ashley, Rodrigues, Michelle, Sokolovic, Nina, Borairi, Sahar, Janmohamed, Zeenat, Jenkins, Jennifer M, & Perlman, Michal. (2024). Can a brief professional development improve early childhood educators' responsivity and interaction quality in child care centers? A cluster randomized controlled trial. *Early Education and Development*, 35(5), 1014-1031.
- Burrows, Tylor. (2024). *Educational Development Leadership: A Distributed Leadership Case Study of a University Teaching and Learning Centre*. Université d'Ottawa| University of Ottawa.
- Carvalho, Luisa, Almeida, Dora, Loures, Ana, Ferreira, Paulo, & Rebola, Fernando. (2024). Quality education for all: A fuzzy set analysis of sustainable development goal compliance. *Sustainability*, 16(12), 5218.
- Dierendonck, Christophe, Poncelet, Débora, & Tinnes-Vigne, Mélanie. (2024). Why teachers do (or do not) implement recommended teaching practices? An application of the theory of planned behavior. *Frontiers in Psychology*, 15, 1269954.
- Egert, Franziska, Fukkink, Ruben G, & Eckhardt, Andrea G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of educational research*, 88(3), 401-433.
- Ferguson, Sarah, & Sutphin, Latanya. (2022). Analyzing the impact on teacher preparedness as a result of using Mursion as a risk-free microteaching experience for pre-service teachers. *Journal of Educational Technology Systems*, 50(4), 432-447.
- GARBA, Lawal Sani, & ALIYU, Mustapha. (2025). Impact of Micro-Teaching Skills on Performance in Classroom Control Strategy among Colleges of Education Students in Northwest, Nigeria. *Federal University Gusau Faculty of Education Journal*, 5(4), 219-224.
- Gülsün, İpek, Malinen, Olli-Pekka, Yada, Akie, & Savolainen, Hannu. (2026). Applying the theory of planned behaviour to examine teachers' intentions to teach in inclusive classrooms and their inclusive practices. *British Educational Research Journal*, 52(1), 538-563.
- Hagger, Martin S, Cheung, Mike W-L, Ajzen, Icek, & Hamilton, Kyra. (2022). Perceived behavioral control moderating effects in the theory of planned behavior: A meta-analysis. *Health Psychology*, 41(2), 155.
- Hamilton, Catherine Adele. (2019). *Exploration of Child Adoption Through the Application of Social Network Analysis*. University of Florida.
- Hou, Mingdi, Lin, Yigang, Shen, Yafei, & Zhou, Hui. (2022). Explaining pre-service teachers' intentions to use technology-enabled learning: An extended model of the theory of planned behavior. *Frontiers in Psychology*, 13, 900806.
- Jabbari, Jason, Chun, Yung, Johnson Jr, Odis, Grinstein-Weiss, Michal, & McDermott, Laura. (2022). Perceptions of school quality and student learning during the pandemic: Exploring the role of students, families, schools, and neighborhoods. *Socius*, 8, 23780231221142955.
- Kılıç, Zülküf, & Özgen, Kemal. (2025). Investigation of Middle School Students' Opinions and Self-Efficacy Beliefs on Mathematical Connection with Using Modelling Tasks. *Journal of Theoretical Educational Sciences*, 18(1), 21-53.

- Kroeger, Stephen D, Doyle, Kathryn, Carnahan, Christina, & Benson, Andrew G. (2024). Microteaching: An opportunity for meaningful professional development. *TEACHING Exceptional Children*, 56(6), 462-471.
- Lee, Jung, Cerreto, Frank A, & Lee, Jihyun. (2010). Theory of planned behavior and teachers' decisions regarding use of educational technology. *Journal of Educational Technology & Society*, 13(1), 152-164.
- Liu, Bojie, Yang, Yu, & Xu, Guandong. (2025). *GenAI-Empowered Virtual Micro-teaching Training System for Preservice Teachers*. Paper presented at the International Conference on Behavioural and Social Computing.
- McDiarmid, G Williamson, & Zhao, Yong. (2023). Time to rethink: Educating for a technology-transformed world. *ECNU Review of Education*, 6(2), 189-214.
- Nigate, Demekash Asregid, Mihiretie, Dawit Mekonnen, & Kassa, Solomon Areaya. (2023). Mentoring during School Practicum: Mentor-Mentee Relationship, Roles Assumed, and Focus of Feedback. *International Journal of Work-Integrated Learning*, 24(4), 491-503.
- Raza, Muhammad Aqeel, Nasim, Farhat, & Javed, Muhammad Latif. (2022). Academic Qualification and Teaching Competences; A Comparative Study of Primary School Teachers. *Pakistan Journal of Humanities and Social Sciences*, 10(1), 234-240.
- Rhodes, Ryan E, & Courneya, Kerry S. (2003). Investigating multiple components of attitude, subjective norm, and perceived control: An examination of the theory of planned behaviour in the exercise domain. *British journal of social psychology*, 42(1), 129-146.
- Shehadeh, Hazem Khaled. (2024). In the Ministry of Education/Jordan. *Business analytical capabilities and artificial intelligence-enabled analytics: Applications and challenges in the digital era*, 2, 207-222.
- Siraj, Iram, Melhuish, Edward, Howard, Steven J, Neilsen-Hewett, Cathrine, Kingston, Denise, De Rosnay, Marc, . . . Luu, Betty. (2023). Improving quality of teaching and child development: A randomised controlled trial of the leadership for learning intervention in preschools. *Frontiers in Psychology*, 13, 1092284.
- Suryana, Dadan, Tika, Rahma, & Wardani, Eka Kusuma. (2022). *Management of creative early childhood education environment in increasing golden age creativity*. Paper presented at the 6th International Conference of Early Childhood Education (ICECE-6 2021).
- Thenga, Muofhe. (2026). The effects of microteaching on preservice teachers' self-efficacy in teaching mapwork. *Journal of Geography Education in Africa*, 9, 91-103.
- Thompson, Carise M, Elliott, Susan J, Meyer, Samantha, Leatherdale, Scott T, & Majowicz, Shannon E. (2024). Attitudes, subjective norms and perceived behavioural control factors influencing Canadian secondary school students' milk and milk alternatives consumption. *Public Health Nutrition*, 27(1), e92.
- Tinnes-Vigne, Mélanie, Houssemand, Claude, Guay, Frédéric, Poncelet, Débora, & Dierendonck, Christophe. (2025). Understanding the use of social and emotional learning in elementary schools: A theory of planned behaviour perspective. *European Journal of Investigation in Health, Psychology and Education*, 15(4), 48.
- Wangchuk, Sangay. (2019). Effects of microteaching on the pre-service teachers' teaching competence—A case in Bhutan. *International journal of education and development using information and communication technology*, 15(1), 133-146.
- Zhang, Jiahua, Pan, Qianqian, Zhang, Di, Meng, Bin, & Hwang, Gwo-Jen. (2024). Effects of virtual reality based microteaching training on pre-service teachers' teaching skills from a multi-dimensional perspective. *Journal of Educational Computing Research*, 62(3), 655-683.

Navigating Collaboration and Internationalization at Universitas Negeri Medan in the 21st Century

Erond Litno Damanik¹ , Suci Pujiastuti², Maya Oktora²

¹ Faculty of Social Sciences, Universitas Negeri Medan, North Sumatra, Indonesia

² Faculty of Languages and Art, Universitas Negeri Medan, North Sumatra, Indonesia

Correspondence: Erond Litno Damanik, Faculty of Social Sciences, Universitas Negeri Medan, North Sumatera Province, Indonesia. E-mail: eronddamanik@unimed.ac.id.
ORCID ID: <https://orcid.org/0000-0003-0408-6852>

Abstract

Internationalization and strategic collaboration have shifted from optional activities to essential tools for higher education institutions, aiming to maintain institutional relevance. Universitas Negeri Medan faces significant challenges in reconciling the traditional mandate as a former Teacher Training and Education Institute with the demands of competitive globalization. Therefore, this research aimed to analyze the dynamics, evaluate achievements, and map obstacles in collaborative governance at Universitas Negeri Medan, using data from 2020 to 2025. The results showed that based on internal data from the Office of International Affairs throughout the period, the university established 46 international Memoranda of Understanding and 112 Memoranda of Agreement, as well as 284 national MoUs and 742 MoAs. These partnerships contribute directly and linearly to the fulfillment of Key Performance Indicators, especially IKU-2, -3, -5, and -6. However, a critical analysis identified structural gaps, including bureaucratic obstacles, limited stimulus funds, and language barriers, that contribute to the low conversion ratio of macro MoUs into active MoAs at the research program level. These results contribute to an evidence-based policy recommendation model to minimize inequality. The evaluation also serves as a strategic basis for government agency performance reports, accelerating the quantum leap toward World-Class University status without abandoning local cultural wisdom.

Keywords: Internationalization, Collaboration, IKU, MoU-MoA, Universitas Negeri Medan

1. Introduction

The global higher education landscape in the 21st century is experiencing a structural transformation driven by globalization, digital interconnectivity, and the demands of integrating Society 5.0. Knowledge, innovation, and human talent are moving transnationally at an unprecedented pace. Universities can no longer function as ivory towers serving only the domestic market. Modern universities are required to transform into global innovation hubs, producing graduates with a global mindset and research with a significant impact on global humanitarian issues (Brilyanti, 2025). In a macro context, strategic collaboration and internationalization have shifted to become absolute necessities and core strategies for maintaining institutional relevance and competitiveness. A World Class University (WCU) is increasingly being defined by the breadth of the global partnership network, the mobility of ideas, and international recognition of the entire academic community. This phenomenon creates both pressure

and significant opportunities for Universitas Negeri Medan to reformulate the institutional development roadmap and remain relevant in an increasingly competitive global environment.

Universitas Negeri Medan, one of the leading state universities in western Indonesia, has a strategically important geopolitical position. This university, located in North Sumatra Province along the Strait of Malacca, is geographically close to vibrant Southeast Asian countries such as Malaysia, Singapore, and Thailand. The unique geographical traits provide Universitas Negeri Medan with a natural advantage as a gateway to the internationalization of education in this part of the archipelago. Moreover, the university has a rich historical legacy as the former Teacher Training and Education Institute (*Institut Keguruan dan Ilmu Pendidikan* [IKIP]). Although the mission extends beyond the production of non-teaching professionals in the 21st century, the university remains a crucial institution for training teachers. A fundamental challenge is how to reconcile the traditional role in teacher education with the demands of globalization. The educators and professionals being produced must be equipped to thrive in multicultural environments and embrace digital technologies across borders. Therefore, effective management of collaboration and internationalization is essential for the future success of Universitas Negeri Medan.

The acceleration of university partnerships during the first half of this decade is reflected in the growth chart of legal documents. Each page of the partnership agreement, especially the Memorandum of Agreement (MoA), is designed to drive progress toward the Key Performance Indicators for State Universities (*Indikator Kinerja Utama-Perguruan Tinggi* [IKU-PTN]). First, IKU-2 focuses on students gaining experience outside the campus. National MoAs cover the implementation of internship and teaching assistant programs in various educational institutions and industries in North Sumatra. Additionally, the international MoAs support credit transfer programs and student participation in the Indonesian International Student Mobility Awards (IISMA). Second, IKU-3 emphasizes lecturer engagement in activities outside the campus. The partnership document provides a legal framework for visiting and guest lecturers at partner universities abroad, or to work as practitioners and experts within government institutions, including the business and industrial sectors.

The research examined two further indicators. Third, IKU-5 pertains to how faculty can achieve international recognition through collaborative research initiatives. These partnerships lead to joint research and efforts that result in publications with leading global researchers in esteemed journals indexed by Scopus or Web of Science (WoS). The collaboration is essential, as it enhances global citation metrics. Fourth, IKU-6 emphasizes the importance of research programs developed in partnership with prestigious international institutions. By forming strategic alliances with top-ranked universities in the QS100-500 range, the curriculum can be updated according to Outcome-Based Education (OBE) standards. This restructuring aims to achieve international accreditation, including recognition from the Accreditation Board for Engineering and Technology (ABET), the Indonesian Accreditation Board for Engineering Education (IABEE), and the Accreditation Agency for Degree Programs in Engineering, Computer Science, Natural Sciences, and Mathematics (*Akkreditierungsagentur für Studiengänge der Ingenieurwissenschaften, der Informatik, der Naturwissenschaften und der Mathematik* [ASIIN]), along with the ASEAN University Network-Quality Assurance (AUN-QA).

Despite having a strong commitment through a Memorandum of Understanding (MoU), Universitas Negeri Medan is encountering several challenges in translating this commitment into an active Memorandum of Agreement (MoA). Several MoUs tend to remain at the ceremonial level without any actual follow-through at the program level. Some obstacles contribute to this situation. First, communication and bureaucratic gaps often limit coordination from the rector's office, specifically the Office of International Affairs (*Kantor Urusan Internasional* [KUI]) and the cooperation division, to the implementing units such as faculties and specific programs. Consequently, administrators may not fully grasp the scope of the MoUs that can be acted upon. Second, limited resources and funding for initiatives pose a significant barrier because actual activities, such as joint research or faculty exchanges, require substantial financial backing. The lack of sufficient internal funding makes it difficult for faculty members to independently pursue international collaboration. Third, language barriers and varying levels of readiness within the academic culture impede effective communication with international partners, thereby slowing the transformation of agreements into concrete collaborative programs.

This research has practical significance as an analytical foundation for preparing the performance report of Universitas Negeri Medan. As an institution operating under the Financial Management of Public Service Agency (*Pengelolaan Keuangan-Badan Layanan Umum* [PK-BLU]), directly accountable to the Ministry, Universitas Negeri Medan must demonstrate performance in an accountable, transparent, and evidence-based manner. This research underscores the importance of assessing partnership achievements not only through quantitative metrics but also by evaluating effectiveness in enhancing the quality of the “Three Pillars” of Indonesian Higher Education (*Tri Dharma Perguruan Tinggi Indonesia*): education and teaching, research and development, and community services. Analyzing the challenges in transitioning from MoUs to MoAs is a critical part of this gap analysis. The results provide recommendations to improve governance strategies in the upcoming year, as Universitas Negeri Medan strives to advance towards WCU status in the 21st century.

2. Theoretical framework

The theoretical framework serves as a conceptual foundation for dissecting, analyzing, and evaluating the dynamics of collaboration and internationalization in higher education, particularly at Universitas Negeri Medan. To build a robust analysis, this research integrates a triple-lens framework: (1) comprehensive internationalization, (2) future institutionalism in education, and (3) the resource-based view (RBV). These three theoretical perspectives are combined through the Quadruple Helix model to address the challenges of the collaboration and internationalization ecosystem in the 21st century. In this research, internationalization is conceptualized as a deeply rooted institutional commitment to embedding international, comparative, and global dimensions into the ethos, mission, and all functions of the Three Pillars of Higher Education. This approach is divided into two operational pillars, namely (1) internationalization abroad, which refers to the physical movement of students, faculty, research projects, and academic programs across national borders, and (2) internationalization at home (IaH), focusing on creating a multicultural campus ecosystem, integrating global perspectives into outcome-based education curricula, and using technology for virtual exchanges of all academics who do not participate in physical mobility (Hudzik, 2011; Soule, 2025).

The first theoretical approach is Hudzik’s (2011) concept of comprehensive internationalization, supported by the results from the sixth global survey of the International Association of Universities (IAU), which gathered data from 722 higher education institutions across 110 countries (IAU, 2024). Macro data show that 77% of universities worldwide considered internationalization to be of great importance (IAU, 2024). Interestingly, the main driver of this shift is not only commercial interests, such as recruiting high-paying international students, but rather the strategic necessity to connect globally with partner institutions and enhance institutional quality (IAU, 2024). From this perspective, Universitas Negeri Medan’s initiative to establish 46 international MoUs is a well-organized response to the evolving landscape of global interconnectivity.

The second theoretical approach is future-oriented educational institutionalism, which examines how higher education institutions adapt to address the complexities of the 21st century (Gidley, Hampson, Wheeler, and Bereded-Samuel, 2010a; Inayatullah and Gidley, 2000). The connection between theory and practice is demonstrated through institutional restructuring and shifts in educational paradigms. The current formal education model reflects the rigid, mechanistic structures of the industrial age, and should be replaced with postformal education that focuses on the evolution of consciousness, transdisciplinary curricula, and human-centered futures (Inayatullah and Gidley, 2000). In the sociology of education, institutionalism views the campus not merely as a learning environment, but an institution shaped by rules, norms, values, and social expectations (Gidley, Hampson, Wheeler, and Bereded-Samuel, 2010b). As a former Teacher Training Institution (*Lembaga Pendidikan Tenaga Kependidikan* [LPTK]), Universitas Negeri Medan has undertaken institutional reforms to remain relevant in an increasingly dynamic educational environment. Incorporating institutionalism theory can transform the institution from simply reproducing past knowledge into one that actively shapes the future. This transformation is essential for ensuring that graduates develop a forward-thinking mindset and are equipped to educate the new generation of the 21st century.

The third theoretical approach is the RBV, which is valuable for analyzing partnership governance and explaining how internal resources are managed to create sustainable competitive advantage (Barney, 1991; 2001). According

to RBV, resources that generate competitive advantage must meet four criteria of being valuable, rare, inimitable, and non-substitutable (VRIN). Within this framework, Universitas Negeri Medan does not aim to simply replicate the Western research university model but rather to leverage its unique strengths. As an LPTK, the university has comparative advantages in educational methodology, curriculum development, and character building. The slogan "*The Character-Building University*" is both a marketing phrase and a strategic resource from the RBV perspective. In the internationalization efforts, the university offers a character-based education model that is functional and infused with the local wisdom of North Sumatra. Furthermore, the RBV framework helps illustrate how Universitas Negeri Medan manages and transforms internal resources into significant bargaining power during international collaboration negotiations. In terms of partnership management, cooperation documents are categorized into two asset types. The first is MoUs, which represent potential resources with macro-legal value, but do not meet the VRIN criteria when unimplemented as "sleeping" MoUs. The second is MoAs, which refer to realized capabilities, such as those implementing a Scopus-indexed joint research scheme, an international class, or credit transfer. These initiatives subsequently evolve into dynamic capabilities that enhance the institution, IKU.

The three theoretical approaches are integrated into the Quadruple Helix model to address the collaboration and internationalization ecosystem in the 21st century. Collaboration in this century is no longer linear or limited to universities (Carayannis and Campbell, 2009). Within the framework of accelerating internationalization, Universitas Negeri Medan applies the dynamics of the Quadruple Helix model, stating that innovation and institutional progress arise from the synergistic interaction among four main actors, including academics, government at both the ministerial and regional levels, as well as industry or the private sector, and civil society (Carayannis and Campbell, 2009; 2012). In this theoretical narrative, the Quadruple Helix becomes a catalyst for strategic collaboration and internationalization programs, reflected as a mutually beneficial ecosystem network in which the university serves as a global knowledge hub for regional progress.

Global data indicate a significant structural gap in internationalization management. Research found that only 11% of public higher education systems worldwide have formal strategic plans and effective operational implementation for internationalization (Huang, Crăciun, and de Wit, 2022). Most successful internationalization resource absorption remains dominated by member countries of the Organization for Economic Cooperation and Development (OECD) (Crăciun, 2018). This pattern is reflected in internal data from Universitas Negeri Medan from 2020 to 2025, showing that 330 MoUs documents (46 international and 284 national) were reduced to 854 active MoAs (112 international and 742 national). From an RBV perspective, the conversion disparity stems from internal constraints such as limited research funding, language barriers, and campus bureaucracy that have not fully adapted to the culture of international collaboration.

This research integrates four conceptual perspectives to map the strategic direction of Universitas Negeri Medan as it navigates collaboration and internationalization in the 21st century. The core of the transformation is driven by the comprehensive internationalization paradigm, emphasizing that internationalization is not only a supplementary program but a comprehensive institutional commitment permeating all aspects of the Three Pillars of Higher Education, from governance to campus culture. The acceleration of this global vision is contextualized through the perspective of future educational institutionalism, which positions Universitas Negeri Medan as an institution adaptive to the disruptions of the times, capable of reforming rigid bureaucratic structures to anticipate future educational trends based on digitalization and curriculum flexibility. To compete internationally without losing identity, Universitas Negeri Medan implements the RBV strategy to identify, manage, and capitalize on valuable and unique internal assets that are transformed into competitive advantages in the global market. The governance of these resources is actualized through the Quadruple Helix ecosystem that integrates the active roles of academia, government, industry, and society in a synergistic innovation network. Through this synthesis, internationalization is understood as a complete strategic movement driven by a comprehensive commitment, oriented towards institutionalism, based on internal character strengths (RBV), and manifested through the Quadruple Helix.

3. Method

Navigating collaboration and internationalization at Universitas Negeri Medan in the 21st century requires an approach that not only captures surface data but also delves into the realities underlying the processes. Therefore, this research applies a mixed-methods approach, a scientific work procedure that combines quantitative and qualitative methods in a single investigation (Creswell and Creswell, 2018). The decision to use mixed methods is based on the argument that a single method is not sufficient to capture the complexity of the dynamics related to higher education internationalization. By integrating these two methods, the inherent biases of each approach can be minimized, resulting in a more valid, objective analysis with a high level of applicability for policymakers. This research specifically adopted an explanatory sequential design (Creswell and Creswell, 2018), which operates in two separate but mutually reinforcing chronological phases. The essence of this design is to use qualitative data to clarify, deepen, and contextualize the quantitative results obtained in the initial phase. In the first phase, the research started with a quantitative approach focused on macro-mapping the extent to which comprehensive internationalization indicators and Quadruple Helix ecosystem involvement have been realized at Universitas Negeri Medan. Numerical data were collected through the distribution of structured Likert-scale questionnaires to a representative sample of respondents, including program leaders, lecturers, KUI staff, and students involved in global mobility programs.

The collected quantitative data were then analyzed using descriptive statistics to determine the percentage of achievement and using inferential statistics to examine structural relationships among variables, such as how institutional governance capacity influences the success of international partnership programs. Table 1 shows a matrix of empirical indicators that integrates the four theoretical foundations into a research instrument for the context of Universitas Negeri Medan.

Table 1: Empirical matrix of research instruments

No	Theoretical Basis	Dimensional Variables	Indicators	Questionnaire items	Interview Focus
1	Comprehensive internationalization	Curriculum and Global Learning	Integration of a global perspective into the syllabus	1,2,3	Curriculum alignment strategies.
		Academic Mobility	International or bilingual classes	4,5,6	Barriers to internationalization abroad and the solutions.
		Campus Policy and Culture	Relevance to international accreditation	7,8	Internalizing a global vision within the work culture.
2	Institutionalism in the Future of Education	Regulatory Flexibility and Adaptability	Campus regulations governing the conversion of international courses, including the speed of bureaucratic responses to global partnerships.	9, 10	The process of transforming the governance of State Universities into Legal Entities and simplifying bureaucracy.
		Digitalization and Future Infrastructure	Readiness of a globally standardized learning management system (LMS) platform and the accessibility of supporting digital infrastructure (virtual mobility).	11, 12	Digital infrastructure readiness to support the hybrid education trend.
3	RBV	Uniqueness of Organized Identity	"The Characters Building University" has a partnership advantage, including unique teaching methodologies.	13, 14	How to package the excellence of character education to attract foreign university partners.

		Capitalization of Local Wisdom	Integration of multiculturalism into global research products and recognition of locality by the international academic community.	15, 16	Optimizing local wisdom assets as a differentiator from other research universities.
4	Quadruple Helix Ecosystem	Synergy Between Academics and Government	Funding and regulatory support for international programs and contributions of expertise to regional development policies.	17, 18	Effective use of government grants to promote international recognition.
		Partnerships with Industry and the Community	Involvement of multinational companies in curriculum design and internships, as well as the downstreaming of international research to address civil society issues.	19, 20	The challenge of aligning global industry expectations with academic research achievements at Universitas Negeri Medan.

The quantitative analysis results obtained from the first phase serve as a bridge to the second phase, namely the qualitative approach. The numbers and trends discovered in the first phase often leave open the questions of "why" and "how." This is where qualitative methods explain the underlying mechanisms of the numerical data. Using a purposive sampling technique, researchers identified key informants, including the Rector, the Vice Rector for Cooperation, the Head of KUI, and foreign industry partners, to participate in in-depth interviews and focus group discussions (FGDs). The qualitative process uncovered the realities behind the numbers, including the psychological and cultural barriers lecturers face when teaching international classes, the bureaucratic challenges in converting course grades, as well as strategies for packaging the identity of "The Character Buildings University" to appeal to global partners.

The final phase of this method is the integration of both types of data in the analysis and discussion. Researchers juxtapose quantitative data visualizations, such as tables or diagrams, with narrative excerpts from qualitative interviews. Therefore, the final interpretation not only presents the status of internationalization achievements statistically but also offers a tactical roadmap and applicable solutions rooted in the real conditions of organizational culture in the 21st century. To ensure the quality, validity, and accuracy of the results of this mixed-methods research with an explanatory sequential design, data validation techniques are divided into three main stages, namely quantitative, qualitative, and integration.

Quantitative data validation emphasized subjecting the instrument to a series of scientific trials to ensure accuracy and consistency, including content and construct validity, as well as reliability. Meanwhile, qualitative data validation, designed to ensure the validity of verbal and textual data obtained from interviews and focus group discussions, emphasizes source triangulation, technical triangulation, member checking, external validity, and reliability (Guba and Lincoln, 1994). Finally, mixed methods validity ensures that the selection of qualitative informants in the second phase is truly based on the statistical results of the first phase, such as selecting research programs with the highest and lowest internationalization achievements based on questionnaire scores and sample consistency. It ensures that qualitative informants are part of or an extension of the quantitative population in the first phase to avoid bias due to differences in case backgrounds. This research was conducted from 2024 to 2025, with a relatively long duration to strengthen the validity of the data in formulating collaboration and internationalization policies at Universitas Negeri Medan.

4. Results

Based on a summary of internal data from the KUI and the Cooperation Section at the Office of the Vice Rector for Cooperation, Public Relations, and Information Systems at Universitas Negeri Medan, a collection of MoUs and MoAs partnership documents signed between 2020 and 2025 is presented. Table 1 shows statistics on legal documents for collaborations between 2020 and 2025.

Table 1: Statistics of legal documents for collaboration with Universitas Negeri Medan, 2020-2025

Region and Document Type	Active Cumulative, 2020-2025	Related Partnership Clusters/Related Institutions
International MoU	46 Documents	Universities in the ASEAN region (Malaysia, Thailand), Taiwan, Pukyong National University (Korea), Australia, France, Germany, and Arizona State University.
International MoA	112 Documents	Research consortia, credit transfer programs, and transnational lecturer exchanges.
National MoU	284 Documents	Ministries, Regional Governments in North Sumatra, State-Owned Enterprises, and Non-Governmental Organizations.
National MoA	742 Documents	Creative Industry Partners, MBKM (Freedom to Learn - Independent Campus) Partner Schools, and Teacher Training Centers.

Source: Office of the Vice Rector for Cooperation, Public Relations, and Information Systems, Universitas Negeri Medan, 2026.

Historical data from 2020 to 2025 show a very positive trend, with Universitas Negeri Medan not only increasing the number of MoUs but also encouraging a rise in the number of MoAs, indicating a real commitment to action at the faculty and program levels. Furthermore, the data demonstrates the extensive network of partnerships. However, the main challenge is no longer simply increasing the number of signatures on paper, but rather optimizing the use of each agreed-upon collaboration document. Data on active partnerships in the form of MoUs and MoAs at the faculty level show a distribution of portions consistent with the focus of the IKU-6, namely, program partnerships. Table 2 shows the cumulative distribution of active national and international collaboration documents per faculty throughout 2020-2025.

Table 2: Distribution of Collaboration per Faculty, 2020-2025

Faculties	Collaboration		Key strategic partners, domestic and global
	National	International	
Faculty of Languages and Arts	65	14	United States Consulate, German Consulate, French Consulate, Universiti Sains Malaysia, Goethe-Institut, North Sumatra Language Center.
Faculty of Mathematics and Natural Sciences	70	12	Kyushu Institute of Technology, Japan, Kasetsart University, Thailand, and the National University of Malaysia.
Faculty of Economics and Business	85	16	University of Sains Malaysia, National University of Malaysia, PT Bank Sumut, and Financial Services Authority.
Faculty of Engineering	95	8	State Electricity Company, Toyota Astra Motor, Vocational Training Centers, and Vocational High School partners.
Faculty of Education	110	7	Provincial/City Education Offices, Bina Guna Sports College, Pembina Laboratory School.
Faculty of Social Sciences	55	5	National Land Agency, Central Statistics Agency, Regional Museums, General Election Commission, and Culture and Tourism Office.
Faculty of Sports and Health Sciences	50	4	Ministry of Youth and Sports, Indonesian National Sports Committee, Asahan University.

Faculty of Medicine	12	0	Bhayangkara Hospital, Hajj Hospital, Tebing Tinggi Regional Hospital, and Adam Malik Hospital.
---------------------	----	---	--

Source: Office of the Vice Rector for Cooperation, Public Relations, and Information Systems, Universitas Negeri Medan, 2026.

Overall, the partnership data by faculty show that the internationalization agenda has progressed across all academic disciplines. Science and business clusters focus on global commercial innovation, while the humanities cluster emphasizes maintaining performing arts and ethnic literature on the regional stage. Based on the annual trend visualization, the dynamics of national and international collaboration data show a very strong recovery pattern (exponential growth) that divides the timeline into two main phases. The first is the pandemic phase, spanning 2020-2021, which reflects regulatory pressure and retraction. During the pandemic, the collaboration graph reflects the nadir of global macro partnership activity, with the number of active international MoUs decreasing to 12 at the start of the pandemic and further shrinking to 8 partnerships in 2021. The causes were large-scale international travel restrictions (lockdowns), restrictions on physical mobility, and a shift in the university budget focus to fulfill emergency online learning infrastructure. The second is the post-pandemic acceleration phase, 2022-2025, which reflects a surge based on the IKU. In 2022, there was a turning point, and with the reopening of international borders, the graph immediately jumped significantly to 19 partnerships through the initiation of hybrid international mobility programs.

From 2023 to 2024, the number of partnerships rose sharply from 32 to 48. The triggers were national scholarship programs such as IISMA and the aggressive pursuit of IKU by faculties (Languages and Arts through the US Consulate and the Faculty of Economics with Universiti Malaya) to pursue the achievement of IKU. Finally, in 2025, the peak achievement of cumulative international partnerships reached 65, marked by a substantial inflow of foreign research and the hosting of cross-continental scientific media, such as international forums initiated by the Faculty of Languages and Arts. The turning point and peak were inseparable from leadership changes at Universitas Negeri Medan at the end of 2023, where the Vice Chancellor for Cooperation, Public Relations, and Information Systems was entrusted by the Chancellor and Chair of the Senate to Erond L. Damanik, a Professor from the Faculty of Social Sciences.

During the Vice Chancellor's tenure, the university successfully established international partnerships with campuses and institutions such as Pukyong National University, South Korea, Arizona State University, United States Agency for International Development (USAID), Le Havre University, France, and Freie Universität Berlin, Germany, as well as 13 QS 200 campuses in Malaysia. The aggressiveness in pursuing the WCU target facilitated a significant increase in the number of lecturers working at top world universities, supported by prestigious scholarship schemes, including Fulbright and internal institutional funds. Short-term guest lecture programs, filled by young and senior lecturers, were held at international-class forums managed by foreign universities, both bilaterally and within regional consortia, throughout 2023-2025. Figure 1 shows the growth trend in international partnerships from 2020 to 2025.

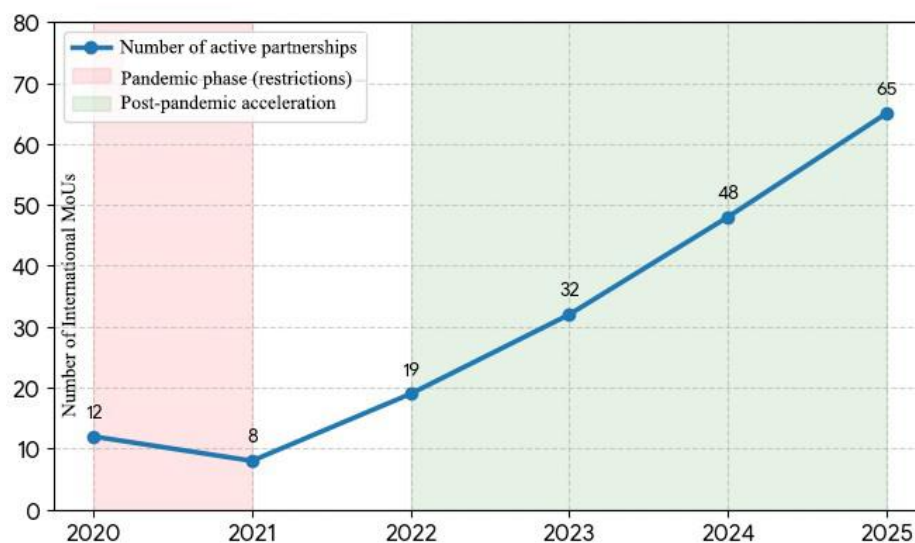


Figure 1: Trends in international collaboration at Universitas Negeri Medan from 2020 to 2025

Source: Office of the Vice Rector for Cooperation, Public Relations, and Information Systems, Universitas Negeri Medan, 2026

Overall, this comparison demonstrates that Universitas Negeri Medan's collaboration and internationalization strategy achieved a full (V-shaped) recovery post-pandemic, transforming the global restriction crisis into opportunities to expand the campus collaborative network in the 21st century. A tangible component of this collaboration is evident in the increase in the Scopus publication index from 2020 to 2025, which shows a strong linear correlation with the expansion of the international collaboration graph. When macro partnership restrictions were lifted post-pandemic, the impact immediately drove the global research output of the academic community exponentially. Table 3 shows a specific comparative analysis that maps the dynamics of Universitas Negeri Medan Scopus publications alongside the previous collaboration growth graph.

Table 3: Correlation of Active MoUs with Scopus-indexed article volume, 2020-2025

Analysis Period	Average International Partnership	Annual Scopus publication volume	Collaboration Characteristics and Research Impact
Pandemic Era (2020-2021)	8-12 Partners	~110-140 articles/year	Dominated by local independent publications or domestic conference proceedings with low citation impact.
Post-Pandemic Era (2022-2023)	19-32 Partners	~220-350 articles/year	The rise of hybrid research in co-authorship formats with researchers in ASEAN.
Excellent Acceleration Era (2024-2025)	48-65 Partners	~480-620 articles/year	The surge is dominated by journals published in Q1–Q4 and by proceedings from globally reputable publications, particularly IEEE.

Source: Office of the Vice Rector for Cooperation, Public Relations, and Information Systems, Universitas Negeri Medan, 2026

The acceleration of the publication graph from 2023 to 2025 was driven by three strategic policy interventions by the rectorate. The first is the strict regulations of the Research and Community Service Institute, where the latest policy requires all International Collaborative Research schemes funded by internal funds to produce mandatory outputs in the form of articles accepted in at least Reputable International Journals, Scopus Quartile 4, or Web of Science (WoS). The second intervention is internationalization of co-authorship, where, through active partnerships with campuses such as Universiti Sains Malaysia (USM) and Kasetsart University, lecturers of Universitas Negeri Medan are involved in joint research schemes. This collaboration breaks the deadlock in penetration into international journals by using laboratory networks and joint funding. The third is a publication-based graduation policy that provides incentives and encouragement for outstanding students to convert the results of final project research into Scopus-indexed international proceedings. Figure 2 is a graph of the cumulative distribution of Scopus articles by quartile throughout 2020-2025.

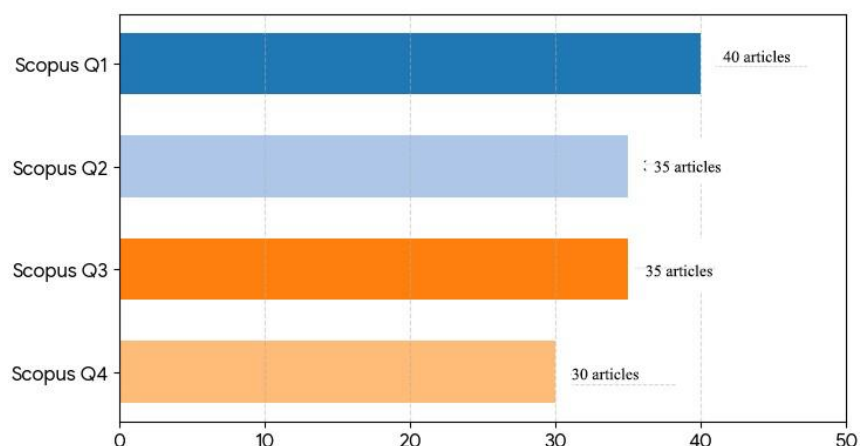


Figure 2: Scopus Articles by Quartile from 2020 to 2025

Source: Office of the Vice Rector for Cooperation, Public Relations, and Information Systems, Universitas Negeri Medan, 2026

The graph above empirically demonstrates that Universitas Negeri Medan's academic transformation in the 21st century no longer relies on the quantity of mass-market articles (non-Quartile) but has instead shifted toward publishing in high-impact journals. This shift is evident in both quantity and quality, as reflected in the Scopus index from 2020 to 2025. The period of 2020-2021 focused solely on publications on conventional, local-scale education topics indexed by Sinta or in non-Scopus proceedings, while 2024–2025 showed a diversification of research clusters toward Artificial Intelligence in Education, gamification, digital-based ethnic culture, and pure science. Co-authorship with international researchers strengthened the university's scientific visibility, marked by a significant increase in citedness within the Scopus database for journals managed internally by the institution, such as the *Journal of Geography* in the Faculty of Social Sciences.

Based on reporting data for the IKU-3, namely lecturers engaged in off-campus activities, as well as records of international partnerships at Universitas Negeri Medan throughout 2020-2025, data on guest lecturers, visiting professors, and visiting lecturers abroad were obtained, showing fluctuations closely correlated with the transition from the pandemic to the post-pandemic acceleration era. Cumulatively, the estimated total teaching assignments and invitations for lecturers to partner universities abroad, both online/hybrid during the pandemic and fully physical post-pandemic, amounted to about 35 to 45 assignments. During the pandemic phase throughout 2020-2021, only around 5-8 assignments were carried out online. Restrictions on global physical mobility (lockdown) shifted teaching activities abroad to digital platforms. Universitas Negeri Medan leveraged regional relations in Southeast Asia to continue sending lecturers to teach in overseas classes virtually. The transition and recovery phase throughout 2022-2023 increased the number of assignments to 12-15, along with the opening of international borders, exchange programs, and the revival of physical teaching. Universitas Negeri Medan lecturers were increasingly invited back to universities in Malaysia, Thailand, and Taiwan as the joint-curriculum program expanded. The global acceleration phase of 2024-2025 increased the number of assignments to 18-22, marking the peak of the institution's performance.

Collaboration-based income generation, both national and international, at Universitas Negeri Medan from 2020 to 2025 is managed centrally through the Public Service Agency (BLU) Non-Tax State Revenue (PNBP) mechanism. As a BLU-status university, Universitas Negeri Medan does not rely solely on student tuition fees (*Uang Kuliah Tunggal* [UKT]) but actively maximizes commercial institutional collaborations that generate non-academic income. Collaboration-based income generation from 2020 to 2021 reached IDR 2.5 billion to IDR 3.5 billion per year, with collaboration income experiencing a drastic decline due to the cancellation of various physical lease contracts, face-to-face training, and domestic industrial field projects. The majority of income was supported by digital project work, independent contract research, and online laboratory testing services of limited value. Post-pandemic income generation from 2022 to 2025 reached IDR 6 billion to IDR 11 billion by the end of 2025. This increase was driven by the commercialization of campus assets, large-scale industrial internship programs, and increased research grants from international institutions.

5. Discussion

The 21st century has fundamentally reshaped the global higher education landscape. Education is no longer limited by geographical boundaries and increasingly demands the fluid exchange of information, technology, and cultural dynamics. For Universitas Negeri Medan, this dynamic is an external challenge as well as a historic call to redefine purpose. Guided by the grand vision of excelling in education, industrial engineering, and culture, Universitas Negeri Medan recognizes that the key to maintaining relevance in this era is grounded in two strategic pillars, namely macro-collaboration and accelerated internationalization. Navigating the institutional direction amid the currents of modernization requires a strong but also adaptive blueprint. Through curriculum transitions, strengthening the quality of human resources, and establishing a network of partnerships across continents, the university is gradually being transformed. From a regional university in North Sumatra, Universitas Negeri Medan has grown into a powerhouse of innovation to be reckoned with on the global stage, steadily moving toward WCU status.

The fundamental steps toward navigating the 21st century must start from the classroom. The rector emphasized the need to produce graduates who not only master theory but also demonstrate high adaptability, creativity, and solution-oriented skills to address the complexities of the modern workplace. A concrete manifestation of this directive is the comprehensive implementation of an Outcome-Based Education (OBE) curriculum across the eight faculties. The OBE approach shifts the focus of education from what lecturers teach to what students are able to achieve upon graduation. To ensure this competency is globally recognized, Universitas Negeri Medan is committed to periodically standardizing the quality of services and programs through international accreditation. This step is complemented by achieving the Excellence Accreditation designation. The recognition demonstrates that the academic quality, curriculum, and assessment system are comparable to those of leading educational institutions worldwide.

As a leading teaching and research institution, Universitas Negeri Medan does not allow academic investigations to be isolated within the campus library. Navigating the 21st century demands high-impact research through international collaboration. This synergy is driven largely by various faculties, particularly through scientific clusters. One remarkable example is the international collaborative research involving the Faculty of Education, Semarang State University, and Kasetsart University, Thailand. The strategic project, funded by the Institute of Research and Community Service (LPPM) through the International Collaborative Research scheme, produced cutting-edge, technology-based instruments and platforms to optimize science and geometry learning at the elementary school level. In the social and economic cluster, the Faculty of Economics is taking aggressive steps by strengthening its international network through a strategic collaboration with Universiti Kebangsaan Malaysia (UKM). The Undergraduate Program in Economics at the Faculty of Economics is also actively implementing International Community Service (*Program Kemitraan Masyarakat* [PKM]). This cross-border collaboration provides opportunities for Universitas Negeri Medan lecturers and researchers to absorb global best practices, exchange perspectives, and publish innovations in internationally reputable journals.

Internationalization extends beyond cooperation documents to the development of an ecosystem that fosters a global mindset across the entire academic community. Universitas Negeri Medan consistently sends students and lecturers to participate in prestigious international mobility programs, including national partnership scholarship programs such as the Indonesian International Student Mobility Awards (IISMA). Through this platform, students from various faculties, including the Faculty of Languages and Arts and the Faculty of Engineering, have the opportunity to study abroad at world-class partner universities. The tangible impact of this global ecosystem is evident in the outcomes of the alumni. Two alumni from the Computer Science Study Program, Angga Warjaya and Sri Mulyana, have successfully passed the international academic selection to continue their master's studies at the Kyushu Institute of Technology in Japan. In addition, the presence of foreign students studying at Universitas Negeri Medan also creates an atmosphere of Internationalization at Home. This daily cultural and academic interaction indirectly trains local students to develop high tolerance, cross-cultural communication skills, and broad worldviews.

Navigating internationalization cannot be robust without strong domestic roots. The university recognizes that global progress must be supported by strong local contributions. This realization has led to consistent expansion of a collaborative network with various universities, local governments, and the domestic industrial sector. For example, the Faculty of Sport Science has established a strategic interdisciplinary partnership with the Faculty of Agriculture at Asahan University (*Universitas Asahan* [UNA]). This unique collaboration brings together expertise in fitness and the agricultural sector to examine crucial issues regarding student mental health and well-being. The collaboration demonstrates the university's ability to view challenges holistically. To produce great scientific and technological innovations in the 21st century, the primary foundation is the psychological well-being and mental readiness of the young generation.

The sustained efforts of faculty at all levels to navigate this collaboration and internationalization have produced significant, concrete results. First, increased credibility for graduates across fields, both educational and non-educational. For example, the Faculty of Social Sciences has demonstrated strong job absorption because graduates are equipped with global competencies. Second, quality management driven by the rectorate and deans has succeeded in converting institutional accreditation (which previously had an A predicate) into a leap in

accreditation of programs with "Excellent" status on the national and international stage. Third, the expansion of international research collaboration has contributed to global publications and citations. Joint research with top campuses in Southeast Asia has increased the number of scientific publications by Universitas Negeri Medan lecturers in reputable international journal indexes such as Scopus.

All achievements in collaboration and internationalization from 2020 to 2025 are inseparable from the leadership of Baharuddin, the Rector of Universitas Negeri Medan, who developed an academic blueprint through a robust theoretical foundation. The concrete implementation of this vision is comprehensively analyzed through a triple-lens framework comprising comprehensive internationalization, the future institutionalism of education, and RBV, all organically integrated into the Quadruple Helix innovation model. This multidimensional synergy has successfully transformed Universitas Negeri Medan from a regional university in North Sumatra into a competitive innovation powerhouse on the global stage, while also demonstrating that global expansion can be achieved without uprooting local cultural identities.

The first perspective, comprehensive internationalization, views global orientation not ordinarily as a ceremonial activity on paper but as a commitment integrated into the entire three pillars of higher education. Concrete manifestations include the shift to an OBE-based curriculum across eight faculties and the attainment of excellent accreditation for various study programs. This global atmosphere is reflected in Internationalization at Home, where students receive digital language testing infrastructure from the Language Technical Implementation Unit (UPT Bahasa) before entering international scholarship programs such as IISMA. The effectiveness of this approach is evidenced in the success of Computer Science alumni from the Faculty of Mathematics and Natural Sciences, Angga Warjaya and Sri Mulyana, who successfully passed the academic selection to continue their master's directly at Kyushu Institute of Technology in Japan.

The second perspective, institutionalism of the future of education, emphasizes Universitas Negeri Medan's agility in adapting to regulatory pressures and global environmental expectations to maintain institutional legitimacy. The surge in international partnerships, from 12 during the pandemic (2020) to 65 active partnerships post-pandemic (2025), represents a strategic response to achieving IKU-6. In this context, the Faculty of Languages and Arts plays a central role by organizing the International Seminar on Language, Arts, and Literature Education (ISLALe 2025). Through this intercontinental scientific platform, the Faculty of Languages and Arts successfully leveraged digital technologies, such as artificial intelligence and gamification, as institutional instruments to redefine the teaching of Malay ethnic literature and arts, ensuring they remain relevant and sought after by Generation Z internationally.

The third perspective, RBV, emphasizes that international competitiveness is largely determined by optimizing valuable and unique internal assets as core competencies. The expertise of senior academics who received Fulbright Visiting Scholar awards in 2024 to the United States or who served as Faculty Ambassadors at Universiti Sains Malaysia (USM) constitutes the primary capital in the joint research scheme. This intellectual excellence is translated into economically valuable products, helping Universitas Negeri Medan rank among the Top 10 Universities with the Most Copyright Applicants in Indonesia in 2025. As a BLU legal entity, research, downstreaming, and asset commercialization can generate income ranging from IDR 9 billion to IDR 11 billion per year. These non-UKT funds are reinvested to finance publication incentives for top-quartile Scopus journals (40 Q1 articles and 35 Q2 articles), creating a sustainable cycle of asset growth.

All the achievements of these three perspectives are consistent with the Quadruple Helix model, which integrates four main pillars. Universitas Negeri Medan, as a higher education institution, serves as a research dynamo and a producer of global talent. Meanwhile, the government provides a regulatory umbrella and strategic reinforcement, such as the collaboration between the Faculty of Sport and Health Sciences and the Ministry of Youth and Sports to develop national sports science, and the Universitas Sumatera Utara sponsorship of the Faculty of Medicine. The industry, as the third pillar, strengthens the university finances through internships and domestic commercial projects, including collaborations with five state-owned banks. Meanwhile, local communities ensure that internationalization remains grounded in the cultural values and traditions of North Sumatra. The interaction

among these four sectors illustrates how the Quadruple Helix model has supported the university's internationalization while preserving regional identity.

The success of Universitas Negeri Medan in navigating collaboration and internationalization in the 21st century is not only a statistical achievement but a structural transformation driven by a fresh academic paradigm. By synthesizing the triple-lens framework and the Quadruple Helix model, the university developed a well-established strategy. This approach formed a conceptual novelty, a fusion of cultural glocalization and modern technology. Universitas Negeri Medan no longer positions North Sumatra local wisdom as a static object of the past but as a high-impact research currency, enabled by artificial intelligence and gamification tools, then disseminated on the world stage through international forums. Furthermore, this strategic novelty is also manifested in a shift in the BLU income-generating model. Universitas Negeri Medan progressively shifted from conventional reliance on leasing physical assets to a financial circularity based on intellectual property, which has empirically been shown to increase the flow of international research grants.

Theoretically, the results contribute to the application of RBV theory in the public sector. Universitas Negeri Medan demonstrates that in the digital era, intangible assets such as the global mindset of the academic community, the expertise of Fulbright Scholar lecturers, and institutional reputation have a greater influence on building long-term competitive advantage than the ownership of physical assets. Furthermore, steps to redefine the Quadruple Helix model suggest post-pandemic higher education institutions are capable of playing the role of central aggregators that simultaneously align the interests of government regulations (IKU achievement), industry profitability (commercial contract research projects), and the real needs of society (preservation of tradition). Practically, the successful transition of the OBE curriculum and the management of non-UKT PNPB funds can serve as a practical blueprint for other LPTK in Indonesia to produce global graduates without losing their reputation as teacher producers. To eliminate the phenomenon of paper partnerships, Universitas Negeri Medan needs to implement a tactical blueprint embedded in a formal, quarterly action plan. This step is designed to break with the tradition of signing ceremonial MoUs and to convert into implementable, measurable MoAs that directly impact the university's IKU achievement. In the first quarter, partnerships identified as passive or not generating tangible activity within 12 months are immediately moved to dormant status. To ensure accountability from the start, each faculty is required to complete a Commitment Worksheet detailing specific IKU achievement targets before the partnership document is processed by the rectorate. In the second quarter, the action plan shifts to the driving phase, which relies on financial incentives and bureaucratic facilitation at the faculty level. Universitas Negeri Medan allocates special stimulus funds from PNPB to support faculty programs that successfully reactivate passive MoUs into operational MoAs. This step is reinforced by establishing the Partnership Clinic (*Kemitraan Klinik*) in each deanery, a regular mentoring forum that trains Heads of Study Programs to compile activity proposals with high bargaining value for external partners.

Systematic acceleration was implemented in the third quarter through the integration of digital information systems and strict field monitoring. Universitas Negeri Medan launched the SIM-KRS (Cooperation Management Information System) Dashboard, a digital platform that tracks the real-time status of document activity and automatically alerts the dean's office when any partnerships are approaching expiration without a signed MoA. The lower-level organizational structure was also strengthened by appointing Partnership Liaison Officers (PLOs) in each faculty. These liaison officers monitor monthly developments in the field and report results directly to the Vice Rector for Cooperation. In this phase, every scientific output from international MoAs is automatically synchronized with the quality assurance system to validate the institution's research performance. The action plan concluded in the fourth quarter with a sustainability phase that prioritizes performance evaluation and full reconciliation. The rectorate implemented a fair Reward and Punishment system, in which the MoU-MoAs conversion ratio is a key component of the Dean's annual performance report. The faculty that records the highest conversion ratio will be awarded an additional budget ceiling for internal research grants in the following fiscal year. This annual cycle concludes with the Universitas Negeri Medan Partners Evaluation Night forum, a space to appreciate Quadruple Helix elements, renew productive cooperation contracts, and firmly sever ties with uncooperative partners. Clean MoAs data and the high economic value collected from this action plan will then serve as fundamental legal documents to accelerate institutional transition into the status of a Legal Entity State University (*Perguruan Tinggi Negeri Berbadan Hukum* [PTN-BH]).

To secure the sustainability of academic and financial performance in the future, this research recommends three strategic policies. First, the university should institutionalize inbound and outbound mentorship schemes, namely securing the stability of the unique intellectual assets by requiring senior academics after overseas assignments to foster a group of young researchers across faculties in a co-authorship cluster to increase the volume of publications in the Upper Quartile Scopus journals (Q1 and Q2). Second, the university should promote collective digitization of cultural assets through integrated clusters, namely, encouraging cross-disciplinary collaboration between faculties to establish a Digital Humanities Innovation Center. Third, the establishment of the Universitas Negeri Medan Industrial Advisory Board through the preparation of regulations that directly involve national and regional-scale industry representatives in the OBE curriculum reconstruction process in each program to ensure a high level of graduate absorption in the workforce. All three provide absolute flexibility to establish a distinct venture capital company, expand global partnerships, and independently recruit foreign talent to accelerate progress toward becoming a WCU.

6. Conclusion

In conclusion, the transformation of Universitas Negeri Medan in the 21st century demonstrates that collaboration and internationalization are not only quantitative measures on paper but depend on the depth of implementation and real impact. By integrating the triple-lens framework, comprising comprehensive internationalization, institutionalism, and RBV, with the Quadruple Helix model, the university has successfully converted its unique internal assets into global competitiveness. The increase from 12 to 65 international partnerships, accompanied by the publication of 75 Scopus Q1 and Q2 articles, indicates the success of the institutional governance. This acceleration is structurally reinforced by regional digital cultural diplomacy and research commercialization, which enhance revenue. To secure the sustainability of this achievement, implementing a formal action plan, enforcing a strict "No MoAs, No MoUs" policy, and institutionalizing the SIM-KRS digital system are decisive steps that must be taken. This orchestration not only enhances the university's achievements but also emphasizes its identity as a resilient institution.

Author Contributions: Conceptualization, Methodology, Investigation, Data Curation, Validation, Formal Analysis, Supervision, Visualization, Erond Litno Damanik; Investigation, Resources, Data Curation, Writing-Original Draft Preparation, Writing-Review & Editing, Suci Pujiastuti; Resources, Data Curation, Writing-Original Draft Preparation, Writing-Review & Editing, Supervision, Project Administration, Maya Oktora.

Funding: This research was funded by Universitas Negeri Medan under grant number 00299/UN33/KPT/2024, and the APC was funded by the Institute for Research and Community Service.

Conflicts of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics approval: All informants in this study have understood the aims and objectives of the study and have given consent for publication.

Data Availability Statement: Not applicable.

Non-AI Acknowledgments: The authors affirm that the entire process of compiling and collecting factual data, formulating theoretical analysis, and writing the editorial was carried out manually and originally by human intellectual capacity, without the use of any generative artificial intelligence tools or models.

Acknowledgements: The authors are grateful to the Rector of Universitas Negeri Medan and all staff, including the Head of the Institute for Research and Community Service, who facilitated this research. In addition, the authors are grateful for the policy support, facilities, and internal funding provided throughout this research.

References

- Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the Resource-Based View. *Journal of Management*, 27(6), 643-650. [https://doi.org/10.1016/S0149-2063\(01\)00115-5](https://doi.org/10.1016/S0149-2063(01)00115-5)
- Barney, J.B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120. <https://doi.org/10.1177/014920639101700>
- Brilyanti, F.A. (2025). Strategy for internationalizing Indonesian higher education institutions to enhance global competitiveness. *The Journal of Indonesian Sustainable Development Planning*, 6(2), 281-296. <https://doi.org/10.46456/jisdep.v6i2.693>
- Carayannis, E., & Campbell, D.F. J. (2009). Mode 3' and 'Quadruple Helix': Toward a 21st century fractal innovation ecosystem. *International Journal of Technology Management*, 46(3/4), 201-234. <https://doi.org/10.1504/IJTM.2009.023374>
- Carayannis, E., & Campbell, D.F. J. (2012). *Mode 3 knowledge production in quadruple innovation helixes: 21st century democracy, innovation, and entrepreneurship for development*. Springer.
- Crăciun, D. (2018). National policies for higher education internationalization: A global comparative perspective. In *European Higher Education area: The impact of past and future policies*, edited by Curaj, A., Deca, L., & Pricopie, R., 57-81. Springer. https://doi.org/10.1007/978-3-319-77407-7_7
- Creswell, J.W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods Approaches*. Sage Publications.
- Gidley, J.M., Hampson, G.P., Wheeler, L., & Bereded-Samuel, E. (2010a.) From access to success: An integrated approach to quality higher education informed by social inclusion theory and practice. *Higher Education Policy*, 23(1), 123-147. <https://doi.org/10.1057/hep.2009.24>
- Gidley, J.M., Hampson, G.P., Wheeler, L., & Bereded-Samuel, E. (2010). Social inclusion: Context, theory and practice. *The Australasian Journal of University-Community Engagement*, 5(1), 6-36.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In *Handbook of Qualitative Research*, edited by N. K. Denzin and Y. S. Lincoln, 105-117. Sage Publication
- Hudzik, J.K. (2011). *Comprehensive internationalization: From concept to action*. Association of International Educators.
- Huang, F., Crăciun, D., & de Wit, H. (2022). Internationalization of higher education in a post-pandemic world: Challenges and responses. *Higher Education Quarterly*, 76, 203-212. <https://doi.org/10.1111/hequ.12392>
- Inayatullah, S., & Gidley, J. M. (2000). *The university in transformation: Global perspectives on the futures of the university*. Bergin and Garvey.
- International Association of Universities. (2024). 6th IAU global survey on: Internationalization of higher education current trends and future scenarios. *Internationalization*. <https://iau.global/internationalization>
- Soule, M. V. (2025). Contextualizing internationalization at home: Academic perspectives from Cypriot higher education in a changing global landscape. *Frontiers Education*. <https://doi.org/10.3389/feduc.2025.1657401>

Practice of Speaking English Language Fluency through Innovating Communicative Tasks

Naomi Velez Borja¹, Arturo Rodriguez Zambrano²

^{1,2} Facultad de Educación, Universidad Laica Eloy Alfaro de Manabí, Manta, Ecuador

E-mail: e0951745405@live.ulead.edu.ec, arturo.rodriguez@uleam.edu.ec
ORCID ID: <https://orcid.org/0009-0009-0359-9462>

Abstract

This study examined the effect of innovative communicative tasks on oral fluency in English among nine children aged 6 to 10 years enrolled at Fundación Submarino Amarillo in Manta, Ecuador. The research used a preexperimental one-group pretest-posttest design during a 35-day intervention. The pedagogical treatment replaced deductive grammar instruction with task-based activities, including information gap tasks, situational role plays, and mystery-solving tasks, all structured to increase Student Talking Time and promote meaningful oral interaction. Data were collected through oral performance assessment before and after the intervention, together with descriptive and correlational analysis of fluency, productivity, articulation, attendance, and formulaic sequences. The results showed improvement across all oral dimensions. Fluency presented the largest gain, followed by productivity and articulation. Individual progress was consistent across the full group, and the correlational analysis indicated a strong positive association between articulation and formulaic sequences, as well as a positive relationship between attendance and posttest fluency. These findings suggest that task-based communicative instruction can strengthen oral performance in young EFL learners when it is implemented in a structured and context-sensitive way. The study supports the use of communicative tasks as a practical strategy for increasing oral participation and reducing passive language learning in early English instruction.

Keywords: Oral Fluency, Task-Based Language Teaching, Communicative Tasks, Student Talking Time, Young EFL Learners, English as a Foreign Language

1. Introduction

Teaching English as a foreign language faces the challenge of translating linguistic knowledge into functional oral performance. In General Basic Education settings, oral fluency still faces limitations due to practices focused on explaining rules and mechanical repetition, which reduces effective student participation time and undermines meaningful interaction in the classroom (Masuram & Sripada, 2020; Richards & Rodgers, 2014).

In organizations that continue to use a grammar-translation method, this problem becomes more important. This method produces passive language competency by emphasizing the memory of structures and restricting opportunities for oral use. In this case, students gather declarative knowledge but find it difficult to negotiate meaning, participate in impromptu conversations, and react adaptably in real-world communication scenarios. The necessity of giving oral skill development top priority through communicative techniques and activities requiring meaningful contact is supported by recent data from Ecuador (Alava Valdéz et al., 2024).

Despite theoretical developments in task-based learning and the communicative approach, there is still a dearth of data unique to General Basic Education students that would enable assessing the impact of interventions centered on Student Talking Time in Ecuadorian situations. From this angle, the current study was carried out during the 2025–2026 school year at the Fundación Submarino Amarillo, which is situated in the Los Gavilanes area of the city of Manta, using a sample of nine pupils between the ages of six and ten. The goal of the project is to improve oral fluency by using creative communicative exercises that boost student talking time and lessen reliance on teacher guidance. The study's focus is solely on oral production; formal grammar or writing are not included.

The significance of the problem lies in the fact that the 6-to-10-year-old stage offers favorable conditions for the acquisition and consolidation of phonological and syntactic patterns, provided that the classroom offers active interaction and frequent opportunities for oral production (Lenneberg & Chomsky, 1967). In the local context, English language training also addresses a social need linked to participation in broader educational and workplace settings. Therefore, this research aims to implement and evaluate innovative communicative tasks designed to maximize effective speaking time and strengthen oral fluency among students at the Fundación Submarino Amarillo.

According to the evaluated literature, teaching grammar alone does not lead to fluency. Task-based learning encourages the smart use of few linguistic resources to accomplish a task and focuses on reaching a communicative objective. Richards and Rodgers (2014) elucidate the benefits of the communicative method to language instruction, and Nunan (2009) centers meaningful interaction around the job. Masuram and Sripada (2020) reveal that task-based materials improve oral abilities.

In contrast to this background, the present proposal adapts the tasks to the playful and technological context of children aged 6 to 10 and assesses changes in speaking time before and after the intervention. In this way, the study provides a context-specific analysis of oral fluency within an Ecuadorian educational setting.

The primary conclusion drawn from the earlier review is that students' oral fluency will be greatly improved by the methodical use of creative communicative tasks. According to a structured observation sheet and a timing register used during the sessions, the secondary hypothesis asserts that the use of technology resources, such as audio recordings and direct timing devices, in conjunction with the instructional materials provided at the institution, will improve students' speaking time. Finding out how the intervention affects oral fluency is the main goal. The quasi-experimental design supports conclusions about the strategy's efficacy and permits a comparison of performance before and after the intervention.

A pre-experimental design that enables the performance of a focused group of nine youngsters (ages six to ten) has been chosen to test these hypotheses. Since measuring STT necessitates a controlled setting where language interactions can be recorded both before and after the intervention, this design is in line with the nature of the issue. We are able to isolate the variable of fluency thanks to the conceptual delimitation to the 2025–2026 period and the exclusion of formal grammar instruction. This ensures that the conclusions drawn are directly related to the efficacy of communicative tasks rather than to general linguistic maturation or theoretical language study.

2. Method

This study employs a pre-experimental design, as it consists of an educational intervention aimed at developing oral fluency in English through innovative communicative tasks. The study spans a period of 35 days and includes

pre- and post-intervention assessments of the participants, with the aim of analyzing changes in their oral expression skills.

2.1 Technics and instruments

The data collection technique was assessment. A pretest and a posttest were administered to establish the baseline and track the evolution of oral performance. The instrument used was the Cambridge English A2 Key, an oral exam administered in two parts in a face-to-face format, with one or two candidates and two examiners. The exam takes eight to ten minutes for each pair of candidates, and its descriptors correspond to the A2 level of the CEFR, which divides language proficiency into six levels ranging from A1 to C2.

2.2 Participant (Subject) Characteristics

The study was conducted with 9 students aged 6 to 10. The gender distribution showed a predominance of females, with 6 students (66.7 percent), while the male group consisted of 3 students (33.3 percent). In terms of educational background, the sample included students from private, public, and charter schools.

Table 1: Descriptive summary of the participants.

Variable	Category	Frecuency (n)	Percentage (%)
Gender	Female	6	66.7%
	Male	3	33.3%
School sector	Private	4	44.4%
	Public	3	33.3%
	Fiscomisional	2	22.2%

To assess the impact of communicative tasks on participants' language development, the following variables and measurement procedures have been defined:

2.2.1. Primary Measures: Oral Fluency

Oral fluency is operationally defined as the student's ability to produce utterances in the target language with a continuous and natural flow, minimizing cognitive interruptions (prolonged pauses or hesitations). This measure will be assessed using an Adapted Fluency Rubric, based on the descriptors of the Common European Framework of Reference (CEFR) for beginner levels (A1-A2). Specific criteria include:

- Temporal Continuity: Reduction of pauses longer than three seconds in simple sentences.
- Capacity to Connect: Connect ideas using simple connectors like and, but, and then.
- Speech Rate: An increase in the number of intelligible words uttered per minute (WPM) during tasks involving free creation.

2.2.2. Secondary Measures: Student Talking Time (STT) and Confidence

As a secondary measure, Student Talking Time (STT) will be quantified, defined as the percentage of time during the class session in which students engage in active language use compared to the teacher's speaking time. These data will be collected through direct timing and analysis of audio recordings of the sessions. Additionally,

Perceived Confidence will be measured, observed through behavioral indicators such as the initiative to initiate exchanges and the reduction of gestures of insecurity during oral expression.

2.2.3. Covariates

To ensure that the results are attributed to the pedagogical intervention, the following covariates have been identified:

- Prior Knowledge: An initial diagnostic test was administered to establish a baseline for vocabulary and grammatical structures.
- Attendance and Exposure: Regular attendance at the foundation will be recorded, given that the frequency of exposure to communicative tasks directly influences the progression of fluency.
- Age Range: Although the group is small ($N=9$), age (6 to 10 years) is considered a factor influencing the complexity of responses and cognitive maturity for negotiating meaning.

2.2.4. Measurement Quality and Transparency

To ensure data reliability, observations will be validated using the Multiple-Point Observation method, with three measurements taken (initial, intermediate, and final). The rubrics have undergone a content validity review to ensure their cultural relevance and suitability for the context of the “Submarino Amarillo” Foundation. It is worth noting that, although preliminary data on “Intrinsic Motivation” were collected using pictographic questionnaires, these are not included in the main analysis of this report in order to maintain a strict focus on oral communication competence.

2.3. Research Design

This study employed a single-group pre-post design. This design was appropriate for the context of the Fundación Submarino Amarillo because it allowed for the observation of changes in oral fluency and Student Talking Time following the implementation of innovative communicative tasks. The absence of a control group limits strong causal inferences; therefore, the results are interpreted based on the intragroup change observed after the intervention.

The task-based approach served as the foundation for the educational intervention. The experimental modification involved substituting communicative activities that required the functional use of language to solve problems for deductive grammar training. Information gaps, situational role-plays, and mystery-solving were the three rotating task types included in the strategy. These assignments encouraged students to actively participate, negotiate meaning, and engage in spoken communication.

Participation was driven by the foundation’s voluntary dynamics and the natural composition of the group. For this reason, non-probabilistic convenience sampling was employed. The 9 participants were not randomly assigned to different conditions, since they were all part of the same pre-existing group. This methodological decision was based on the feasibility of the study in a non-formal education setting.

The Yellow Submarine Foundation carried out the intervention in 2025–2026. Two 60-minute sessions each week were required per the protocol. A 5-minute pre-task phase, a 40-minute task phase, and a 15-minute final feedback phase comprised each session. Previous vocabulary was used in the first stage. The communicative task was completed in the second phase with little assistance from the researcher. Communicative accomplishments and the most common mistakes were examined in the third stage.

2.3.4 Experimental Manipulations or Interventions

The study adopted a preexperimental one-group pretest-posttest design. This choice fit an intervention context in which the same participants received the treatment and later served as their own comparison baseline. Such designs are commonly used to estimate change after an intervention when full experimental control is not feasible.

The intervention was a systematic pedagogical approach that used task-based language instruction to improve oral fluency in English. The approach substituted communicative tasks that required students to solve a problem, exchange meaning, and produce oral language for a predetermined end for deductive grammar teaching. This approach aligns with the fundamental tenets of task-based language instruction, which prioritize meaningful task completion, learner participation, and communication through interaction.

The intervention was implemented in three rotating task types. The first task type was an information-gap activity in pairs, in which each learner received partial information and had to negotiate meaning to complete the shared goal. The second task type was a situational role-play based on everyday contexts adapted to the children's age group. The third task type was a mystery-solving activity in which learners followed oral or visual clues in English to reach a group solution. These task types align with established task-based pedagogical components and with recent classroom evidence that role-play supports communicative competence, fluency, and confidence.

The intervention was conducted at the Fundación Submarino Amarillo between 2025 and 2026. It consisted of two 60-minute sessions per week. There were three stages to every session. During the five-minute pre-task phase, a quick warm-up was used to activate past vocabulary. The researcher served as both a facilitator and an observer during the 40-minute main task phase, which focused on task completion. The 15-minute post-task phase was devoted to reflection, feedback, and fixing common communication mistakes. The traditional task cycle employed in task-based lessons, which divides instruction into pre-task, main task, and post-task stages, is reflected in this order.

The study involved 9 students between 6 and 10 years of age. The sample was selected by convenience because the institution had a natural preexisting group and the program followed a voluntary structure. No random assignment or separate experimental conditions were used. All participants received the same intervention and the same instructional sequence. This arrangement was consistent with the practical limits of a pre-experimental design and with the educational setting in which the intervention was carried out.

No external control group was included. The manipulation consisted of the planned introduction of communicative tasks, which replaced the usual grammar-centered routine during the intervention period. The comparison relied on the change observed within the same group from pretest to posttest. Consequently, the study evaluated the effect of the treatment through intragroup variation in oral performance rather than through comparison with a parallel group.

3. Results

This section presents the findings from the intervention using innovative communicative tasks designed to develop oral fluency in English. The analysis was organized sequentially, with a description of recruitment, statistical synthesis of the data, and a review of individual change. The study employed a single-group pre-post pre-test design, appropriate for examining within-group variation following an educational intervention (Dimitrov & Rumrill, 2003).

3.1. Recruitment

Recruitment took place at the Yellow Submarine Foundation with a pre-existing group of nine students aged 6 to 10. The draft report does not specify the exact dates of recruitment or follow-up; therefore, this information should be included in the final version of the report. The selection was based on institutional availability and voluntary participation in the program.

3.2. Statistics and Data Analysis

The statistical analysis was based on the calculation of pre-test and post-test means, along with the identification of absolute differences between the two measurements. The analysis focused on three dimensions of oral

performance: fluency, productivity, and articulation. Additionally, individual changes were examined to verify the consistency of improvement within the evaluated group.

Table 2: Comparison of pre-test vs. post-test means.

Dimension	Pre-test Mean	Post-test Mean	Increase
Fluency	1.67	2.78	+1.11
Productivity (Output)	1.67	2.44	+0.77
Articulation	1.78	2.44	+0.66

The results show an increase in all three measured dimensions. Oral fluency showed the greatest improvement, with a net difference of 1.11 points. Productivity increased by 0.77 points, and articulation improved by 0.66 points. Overall, the statistics show that oral performance improved after the intervention.

An increasing tendency among participants was confirmed by individual analysis. The most gain was shown by student F9-3, whose fluency improved by two points. F10-6, M9-5, and M8-3 all gained one point in the meantime. The stability of the change shown in the group average is supported by this pattern.

The intervention maintained a task-based approach with an emphasis on oral exchange and the negotiation of meaning, a central principle of task-oriented teaching (Nunan, 2009).

3.3. Ancillary Analyses

The supplementary analysis was limited to Spearman's correlation between attendance, fluency, productivity, articulation, and formulaic sequences. No subgroup analyses or additional adjusted procedures were applied. With a value of 0.58, attendance demonstrated a somewhat positive relationship with posttest fluency. Similarly, there was a strong positive link (coefficient of 0.92) between articulation and the usage of formulaic sequences. Additionally, a substantial correlation (coefficient of 0.82) was found between productivity and fluency. When combined, these findings show a steady correlation between speech performance attained at the end of the intervention, formulaic expression usage, and session attendance.

3.4. Participant Flow

The initial sample consisted of nine students. According to the submitted data, all participants completed the pretest, the intervention, and the posttest. No dropouts, lost to follow-up, or crossovers between conditions were reported, as the study involved a single natural group with no experimental reassignment. The final version of the report should include the exact dates of recruitment, the start of the intervention, and the end of follow-up, along with the institutional source of recruitment.

3.5. Intervention or Manipulation Fidelity

The intervention was implemented according to the planned sequence of innovative communicative tasks. Each session was organized into an initial activation phase, a central task phase, and a final feedback phase. Attendance monitoring and session records allowed for verification of participants' regular exposure to the program. No substantive changes in content or in the planned order of activities are reported. To complete this section, the final report must include the number of sessions conducted compared to the total scheduled and observational evidence supporting adherence to the protocol.

3.6. Baseline data

The oral production pretest was given before the instructional intervention to establish the study's baseline. The preliminary findings consistently showed poor performance in the areas of articulation, productivity, and fluency. Fluency had an initial mean of 1.67, productivity was 1.67, and articulation was 1.78. These findings verified that the group's oral performance limits at the beginning of the trial were comparable.

Nine students between the ages of six and ten made up the demographic distribution. The male group made up 33.3 percent, while the female group made up 66.7 percent. Additionally, the sample's educational diversity was enhanced by the participants' attendance at public, private, and public-private partnership schools.

Table 3: Baseline linguistic performance in the pretest

Variable	M	SD	LL
Fluency	1.67	0.50	1.17
Productivity	1.67	0.50	1.17
Articulation	1.78	0.44	1.34

3.6.1. Statistics and data analysis

The statistical analysis included all participants who completed the intervention process and the final assessment. Since no student dropped out of the study or was absent from the posttest, the analyzed sample matched the initial sample. For this reason, the analysis focused entirely on the nine participants who underwent the intervention.

Descriptive statistics and Spearman's correlation analysis were used in the data analysis. The means from the pretest and posttest could be compared thanks to descriptive measures. The correlation analysis then looked at the relationship between formulaic sequences, attendance, fluency, and articulation. Given the limited sample size and the ordinal nature of some of the data, this method proved suitable.

With a coefficient of 0.58, the findings indicated a somewhat positive relationship between attendance and fluency. Similarly, a strong positive correlation (coefficient of 0.92) was found between articulation and formulaic sequences. Additionally, a positive correlation (coefficient of 0.82) was found between productivity and fluency. These correlations imply that the participants' oral production was enhanced by repeated usage of formulaic sequences and prolonged exposure to communicative activities.

To better understand the learning process, a Spearman correlation analysis was conducted, identifying two main results:

A moderate-to-strong correlation ($r_s = 0.58$) was identified between attendance at the sessions and the fluency achieved on the post-test, indicating that continuous exposure to innovative communicative tasks has a positive influence on progress.

The employment of formulaic sequences and articulation were shown to be significantly correlated ($r_s = 0.92$). This finding suggests that by lowering the cognitive load during oral production, the usage of pre-made expressions enabled students to talk more rapidly and clearly.

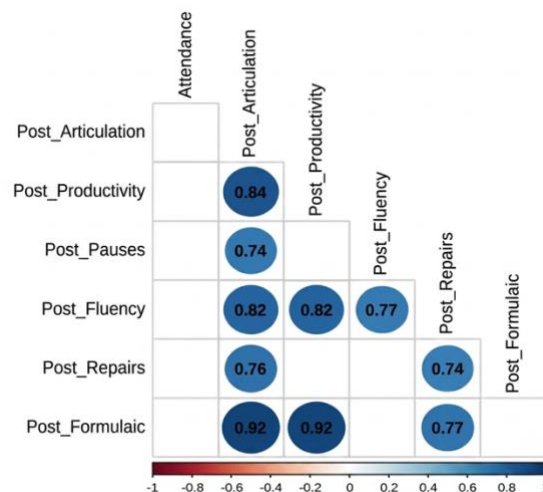


Figure 1: Matrix of Linguistic Correlations

Figure 1: Spearman's correlation coefficients between oral production variables in the post-test.

Source: Prepared by the authors based on data processed in R (2026).

The statistical association between the linguistic parameters measured after the intervention is seen in this image. The most important result is the perfect or almost perfect positive connection ($r_s = 0.92$) between the employment of formulaic sequences and articulation, indicating that using pre-formulated phrases promotes more fluid motor execution. Similarly, there is a considerable correlation ($r_s = 0.82$) between productivity and fluency, indicating that students' perceptions of fluency are inextricably related to speaking at a higher volume.

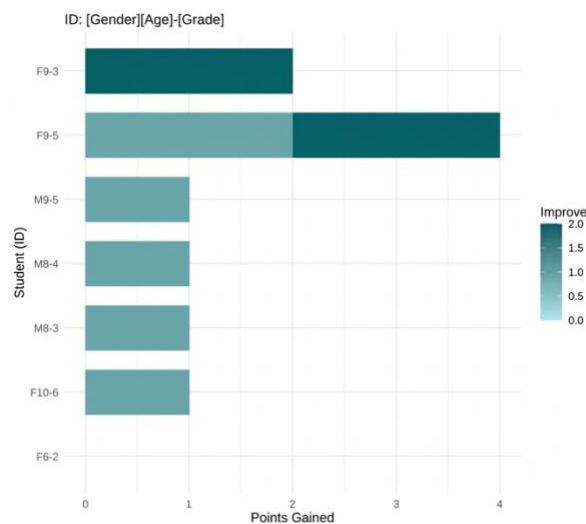


Figure 2: Analysis of Individual Improvement by Student Profile

Figure 2: Distribution of points gained in fluency per participant by gender, age, and grade.

Source: Author's own work (2026)

The nine participants' individual development is shown in the segmented bar chart. Remarkably, every single sample showed an improvement of at least one point on the fluency measure. Student F9-3's 2-point rise is the most noteworthy example. The results show that the creative communicative activities were beneficial to all participants, even the youngest (6 years old) and those who were closer to pre-adolescence (10 years old).

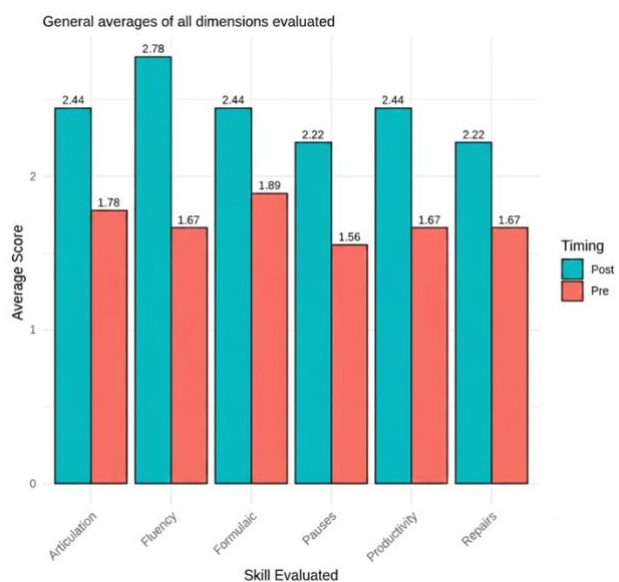


Figure 3: Comparison of Linguistic Performance: Pre-test vs. Post-test

Figure 3: Comparison of overall averages by linguistic dimension before and after the intervention.

Source: Author's own work (2026).

The quantitative effects of the communication tasks on the six main study areas are shown in this image. All variables indicate an increasing trend, but fluency shows the biggest gain, going from an average of 1.67 to 2.78 (+1.11 points). Additionally noteworthy is the rise in Repairs and Pauses, which, in the context of meaning negotiation, suggests that the kids are paying closer attention and exerting more effort when trying to communicate complicated ideas in the target language.

3.6.2. Adverse events

No serious adverse events or side effects that would have compromised the students' participation were reported during the intervention. All participants completed the scheduled activities and maintained regular attendance at the sessions.

However, small episodes of initial insecurity were noted in several activities, particularly in the initial sessions of spontaneous output. Long pauses, a brief drop in speech level, and reluctance to start oral conversations were some of these symptoms. As the intervention went on, these reactions progressively declined and didn't necessitate stopping activities or making significant changes to the teaching methodology.

4. Discussion

The findings support the primary hypothesis, as the intervention led to consistent improvements in fluency, productivity, and articulation, with a notable increase in oral fluency. The secondary hypothesis receives partial support, as the institutional setting—with its educational and technological resources—likely facilitated oral participation, although the study design did not isolate this factor independently. Overall, the observed pattern aligns with recent evidence attributing a favorable effect of task-based language teaching on oral production in school settings, especially when the task organizes verbal exchange with a clear communicative goal.

The most plausible causal interpretation attributes the improvement to the reorganization of the classroom toward greater student production and a reduced centrality of teacher explanation. The increase in Student Talking Time fostered meaning negotiation, turn-taking, and spontaneous production, so that language ceased to operate solely on a receptive level. Recent literature indicates that positive perceptions of teacher speech and classroom interaction are associated with better linguistic performance, while more participatory environments tend to foster greater engagement and learning (Wu & Wang, 2025).

Particularly instructive is the relationship found between articulation and formulaic sequences. Pre-formulated phrases may have acted as a cognitive assist by lowering the need for real-time syntactic building and freeing up resources for oral production, as indicated by the high association. This explanation aligns with Yi y Zhong's (2024) recent meta-analysis, which showed a processing advantage for multi-word sequences, as well as research showing that speed, smoothness, and the effective management of recurrent linguistic units are associated with fluency. Thus, speaking more naturally is just as important to grow as speaking more.

The fact that every participant showed improvement supports the notion that the intervention worked for everyone and wasn't just for a select few. The small group's consistent progress indicates that the tasks' order was appropriate for the participants' age, skill level, and classroom environment. This result is consistent with evidence from Aaj et al. (2024), who discovered that young English learners profited from repeating assignments, especially when it came to oral performance. Accordingly, frequent exposure, controlled variety, and systematic repetition seem to boost confidence and fluency.

The group's positive response also suggests that the instructional sequence was effective because it provided cumulative opportunities for language use, rather than through isolated instruction. Information-sharing tasks, role-playing, and mystery-solving increased the need to respond, clarify, repair, and continue the discourse. This dynamic is consistent with recent studies showing that the task-based approach is most effective when activities are grounded in the real-world classroom context and when oral performance is assessed in situations of meaningful production (Lu et al., 2025). Thus, the intervention did not work through the transmission of rules, but rather through the activation of use.

The study has risks to internal validity, necessitating a cautious interpretation despite the positive outcomes. Regression toward the mean, test familiarity, practice effects, and natural maturation cannot all be fully excluded due to the single-group design. Estimates of change may also be erratic due to the limited sample size and short intervention duration. Any claims for retention are further limited by the lack of a follow-up measurement. As a result, the data indicate that the intervention and performance are positively correlated, but it does not support strong causal conclusions or wide generalizations.

External validity also calls for caution. The study was conducted at a specific institution, with a small group of children aged 6 to 10 and within a unique classroom setting. Therefore, the transferability of the findings to other contexts depends on similarities in developmental level, group size, availability of pedagogical support, and institutional culture. Even so, the research does offer a useful insight for English-as-a-foreign-language settings that need to increase oral use at an early age. The practical relevance is strengthened because recent studies highlight that oral production improves when tasks are tailored to the local context.

The fidelity of the intervention was a decisive factor, as the observed effect depends on how the tasks were carried out. The pre-task, task cycle, and post-task structure allowed for a clear progression, with prior activation, central production, and final feedback. This order aligns with recent approaches to the localized implementation of the task-based approach and with the need for teacher mediation that maintains the communicative focus without reclaiming the monopoly on speech. The literature on teacher talk itself indicates that the quality of teacher-student interaction influences student learning and well-being (Lu et al., 2025).

Alternative explanations should also be considered. Some of the improvement may be due to the combined effect of repeated practice, familiarity with the activity formats, and increased confidence, rather than a single linguistic mechanism. Similarly, the association between fluency and formulaic sequences may reflect a general level of oral competence, not a direct causal relationship between the two variables. The magnitude of the observed correlations, especially with such a small sample, calls for cautious interpretation. At this point, recent evidence on multi-word sequence processing supports the theoretical plausibility of the finding, but does not eliminate the need for replication with larger samples (Yi & Zhong, 2024).

Despite these reservations, the study's significance is clear. In a context where English instruction often prioritizes reception and repetition, the intervention demonstrated that a task-centered pedagogical approach can empower students to speak and transform the classroom into a space for meaningful oral production. This contribution is valuable for teaching practice and for future research, as it suggests that early fluency can be strengthened through contextualized tasks, functional repetition, and interactive mediation. The main contribution lies in the fact that the problem does not seem to lie in a lack of content, but rather in the structure of participation that the classroom offers.

Author Contributions: Conceptualization, methodology, research, data curation, formal analysis, visualization, and original draft writing are among the author's contributions. Project management, writing evaluation and editing, and supervision. Each author acknowledged responsibility for the work's accuracy and integrity and authorized the submitted version.

Funding: No outside funding was obtained for this study. There was no outside sponsor to cover the publication costs.

Conflicts of Interest: No conflicts of interest are disclosed by the authors. The study's design, data collection, analysis, and interpretation, manuscript writing, and decision to publish the findings were all independent of the funding sponsors.

Informed Consent Statement and Ethics Approval: Prior to data collection, each participant's legal guardian provided written informed consent. Additionally, the students' consent was acquired in age-appropriate language. The Declaration of Helsinki was followed in the conduct of the study.

Data Availability Statement: Upon reasonable request, the relevant author will provide the data supporting the study's conclusions. The complete dataset is not publicly stored due to participant age and institutional confidentiality regulations.

Acknowledgments: The Fundación Submarino Amarillo's administrators, instructors, students, and families are acknowledged by the authors for their cooperation and support during the study's progress.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Aaj, A., Maftoon, P., & Siyyari, M. (2024). Do young EFL learners benefit from task repetition? *The Language Learning Journal*, 52(6), 650–661. <https://doi.org/10.1080/09571736.2023.2213706>
- Alava Valdéz, M. J., Cedeño Zambrano, J. R., & Macias Sera, R. (2024). Communicative approach in the development of oral skills in english. *Revista Chone Ciencia y Tecnología*, 2(2). <https://doi.org/10.56124/cct.v2i2.001>
- Dimitrov, D. M., & Rumrill, Jr., P. D. (2003). Pretest-posttest designs and measurement of change. *WORK: A Journal of Prevention, Assessment & Rehabilitation*, 20(2), 159–165. <https://doi.org/10.3233/WOR-2003-00285>
- Lenneberg, E. H., & Chomsky, N. (1967). *Biological foundations of language*. Wiley.
- Lu, J., Ma, Q., & Li, S. (2025). Effect of localized task-based language teaching on Chinese secondary school English learners' oral production in examination-oriented contexts. *International Journal of Applied Linguistics*, 35(1), 168–192. <https://doi.org/10.1111/ijal.12608>
- Masuram, J., & Sripada, P. N. (2020). Developing speaking skills through task-based materials. *Procedia Computer Science*, 172, 60–65. <https://doi.org/10.1016/j.procs.2020.05.009>
- Nunan, D. (2009). *Task-based language teaching* (Reprinted). Cambridge Univ. Press.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and Methods in Language Teaching* (3a ed.). Cambridge University Press. <https://doi.org/10.1017/9781009024532>

- Wu, P., & Wang, Z. (2025). The impact of the student perception of teacher talk on student language learning attainment. *BMC Psychology*, 13(1), 221. <https://doi.org/10.1186/s40359-025-02508-3>
- Yi, W., & Zhong, Y. (2024). The processing advantage of multiword sequences: A meta-analysis. *Studies in Second Language Acquisition*, 46(2), 427–452. <https://doi.org/10.1017/S0272263123000542>

Idea Transformation: Naming the Missing Process in Source-Based Writing

Pauline P. L. Chin¹ 

¹ Meragang Sixth Form College, Brunei Darussalam
ORCID ID: <https://orcid.org/0000-0001-5635-9141>

Abstract

Source-based writing tasks require students to do more than locate relevant information in a text. They require students to convert that information into understood meaning before expressing it in written language. This paper argues that the process connecting these two acts, here termed idea transformation, is an overlooked educational concern that has not previously been examined as a unified teaching objective. The paper draws on classroom evidence from a cohort of twenty-nine Cambridge O Level English Language students participating in a targeted intervention programme at a Sixth Form College in Brunei Darussalam. Two sources of evidence are examined: the cohort's first internal comprehension assessment, which revealed a consistent pattern of retrieval without transformation across twenty-nine scripts, and a single directed writing lesson, in which one student made the missing process visible in a precise and observable moment. The theoretical framework positions idea transformation within established scholarship, drawing on Kintsch's (1988) construction-integration model, Anderson and Krathwohl's (2001) distinction between remembering and understanding, and Bereiter and Scardamalia's (2013) account of knowledge telling and knowledge transforming. The paper argues that students who stop at retrieval are not failing to read or to write. They are missing the intermediate act between the two. The discussion considers what this absence looks like in the classroom, what the existing literature reveals about why it occurs, and what teaching must do if it is to make idea transformation visible and necessary. The paper concludes by proposing that naming this process is the first step towards teaching it deliberately.

Keywords: Idea Transformation, Reading-Writing Relationship, English Language Teaching, Directed Writing, Knowledge Transforming

1. Introduction

Text-based writing tasks are a central feature of the Cambridge O Level English Language examination, and directed writing is among the most demanding of these tasks. Students are required to read two source texts, extract relevant information, and produce a coherent written response in a specified format, typically a letter, report, or email, addressed to a named audience. The task demands more than reading comprehension and more than writing competence. It demands something between the two: the capacity to take an idea from a text and make it one's own before committing it to paper. For many students at the secondary level, this intermediate capacity is precisely what is missing. They can read. They can write. What they cannot consistently do is transform retrieved information into understood meaning before committing that meaning to paper.

This observation emerged from a directed writing lesson conducted during the 2026 O Level English Language Intensive Programme at a Sixth Form College. The cohort comprised twenty-nine students resitting the Cambridge O Level examination, the majority having obtained grades D and E in the October/November 2025 sitting. During one lesson, the researcher observed a female student working on a directed writing task. The student read Text A, identified a relevant sentence, and highlighted it. She then turned to Paragraph 2 of her planning scaffold, which carried the sentence starter: *Firstly, one important point is...* She copied the highlighted sentence directly into her response, with very little change in wording. The student knew precisely where the idea belonged. She placed it correctly. What she did not do was process the idea before placing it.

The observation was not surprising in isolation. What made it significant was that the student had not failed the task by any surface measure. The information she selected was relevant. The paragraph was structurally correct. The sentence starter was in place. She had done everything the scaffold asked of her. And yet the response demonstrated information transfer rather than idea transformation. The idea had moved from the source text to the answer sheet without passing through the student's own understanding.

Assessment data from the cohort's first internal assessment, a school-based comprehension paper, confirmed that this pattern was not confined to one student. Students appeared increasingly capable of locating and retrieving information from texts. The transition from retrieval to interpretation, however, remained inconsistent. Across twenty-nine scripts, teacher commentary returned repeatedly to the same observation: students could identify relevant material, but could not yet explain it, develop it, or express it in language that was demonstrably their own. The challenge was not comprehension in the decoding sense. It was comprehension in the meaning-making sense, the capacity to hold an idea long enough to understand it before writing it down.

The existing literature addresses aspects of this problem from several disciplinary directions, but the specific process connecting these two moments, what happens between locating an idea and writing it down, has not been sufficiently named or examined as a unified teaching concern in source-based English language writing. Bereiter and Scardamalia (2013) distinguished between knowledge telling and knowledge transforming in written composition. Kintsch (1988) separated what a text says from what a reader constructs from it. Flower and Hayes (1981) described writing as a problem-solving process in which meaning is built rather than transcribed. Wittrock (1974/2010) argued that learning occurs only when learners generate meaningful relationships between new information and prior knowledge. Wiggins and McTighe (2005) demonstrated that genuine understanding requires explanation, interpretation, and application. Anderson and Krathwohl (2001) located remembering and understanding as distinct operations.

Bereiter and Scardamalia (2013) and Flower and Hayes (1981) have each examined what the writing process requires of a learner. Kintsch (1988) and Wittrock (1974/2010) have each examined how readers construct meaning from text. Wiggins and McTighe (2005) and Anderson and Krathwohl (2001) have each examined what genuine understanding looks like and what it demands of a learner. What none of these bodies of work has addressed directly is the point that sits between them: the moment in which a learner must stop receiving an idea and begin doing something with it. This paper proposes that this process, observable in classroom practice and supported by existing theory, can be usefully named and examined: idea transformation, defined here as the conversion of retrieved information into understood meaning before that meaning is committed to writing.

Two questions drive the inquiry. The first is conceptual: what is idea transformation, and where does it sit in the sequence that connects reading to writing? The second is practical: what does its absence look like in the classroom, and what does this suggest for teaching? The paper is organised as follows. The theoretical framework section draws on the reading and writing literature to define idea transformation and position it within existing scholarship. The discussion section returns to the classroom evidence, drawn from both the cohort's first internal comprehension assessment and a subsequent directed writing lesson, and examines both in the light of the theoretical framework. The conclusion draws the argument to a close and considers what it means for classroom teaching and for future research.

2. Theoretical Framework

2.1 *Reading Comprehension as Construction*

A reader who locates a relevant sentence in a source text and a reader who understands that sentence have not necessarily done the same thing. This distinction, which lies at the heart of the problem this paper examines, is supported by two bodies of theoretical work: Kintsch's (1988) account of how readers construct meaning from text, and Wittrock's (1974/2010) account of how learners generate meaning from new information.

Kintsch (1988) proposed that reading comprehension involves two related but distinct operations. In the first, a reader builds a representation of what the text says, drawing on the words, sentences, and propositions the text provides. In the second, this representation is integrated with the reader's existing knowledge to produce a fuller, more personalised understanding of what the text means. A reader who completes only the first operation has processed the text at a surface level. A reader who completes both has constructed meaning from it. The difference between these two outcomes is not visible in the text the reader produces immediately after reading. It becomes visible only when the reader is asked to explain, apply, or restate an idea in language that is demonstrably their own.

Wittrock (1974/2010) approached the same territory from an educational direction. Learning, he argued, is not the reception of information but the generation of meaningful relationships between new information and what the learner already knows. A learner who receives new information without generating such relationships has not yet learned in any meaningful sense. The information has been encountered; it has not been made one's own. Wittrock described this generative process as active and deliberate: it requires the learner to engage with new material, to connect it, elaborate it, and integrate it into existing understanding. Passive reception, however attentive, does not produce the same result.

These two accounts establish a principle that is foundational to the argument developed in this paper. Understanding is not a condition that results automatically from reading. It is a process that the reader must perform. The student who reads a source text, identifies a relevant sentence, and copies it into a written response has engaged with the text at the level of retrieval. Students may appear to have read successfully because they can locate a relevant sentence, but location is not the same as comprehension (Tovani, 2000). That student has not yet performed the operation that Kintsch (1988) and Wittrock (1974/2010) each describe as essential: the active construction of meaning from what has been read. What is missing is not attention, and it is not effort in the surface sense. What is missing is the internal process by which retrieved information becomes understood meaning. That process is what this paper sets out to define.

2.2 *The Reading-Writing Relationship*

If reading comprehension is a constructive act, then the move from reading to writing is not a simple handover of content from one activity to another. It is a continuation of the same meaning-making process under new demands. This is the central claim of the reading-writing relationship literature, and it is directly relevant to understanding what idea transformation requires.

Tierney and Pearson (1983) argued that reading and writing are both acts of composing. A reader, like a writer, is not a passive receiver of meaning but an active maker of it. Both reader and writer draw on prior knowledge, set goals, make decisions about what matters, and revise their understanding as they proceed. This framing displaced the older view of reading as reception and writing as production, and proposed instead that the two activities share underlying processes of meaning construction. Shanahan (1988) built on this by demonstrating that reading and writing draw on overlapping bodies of knowledge and that students benefit from instruction that treats them as mutually informing rather than as separate and sequential skills.

What this means for the source-based writing task is significant. A student who reads a source text and then writes from it is not completing two separate activities in sequence. Reading to write is a single, integrated act of meaning-

making in which the knowledge constructed through reading becomes the material from which written meaning is built. Spivey (1990) described this as the constructive work of a reader who is also a writer. Drawing on multiple source texts to produce a new text, the writer performs three operations: organising the material into a structure, selecting what is relevant, and connecting source content with knowledge already held. These are not mechanical operations. Each one requires the writer to make a judgement about meaning. Selecting requires the writer to understand what is relevant and why. Organising requires the writer to see how ideas relate. Connecting requires the writer to bring prior knowledge to bear on new material and to generate something that was not present in the source.

What the existing literature on the reading-writing relationship establishes, then, is that the movement from retrieved information to written response is not a transfer but a transformation. The source text provides the raw material. The writer's task is to remake that material into new meaning. When a student copies a retrieved sentence into a response without processing it, the movement from reading to writing has occurred, but the transformation has not. Retrieval has occurred, and so has writing. What has not occurred is idea transformation: the work that must happen between them.

2.3 Writing from Sources

The literature reviewed in the preceding sections establishes that reading comprehension is a constructive act and that writing from sources requires the reader to become a maker of new meaning. What the source-based writing literature adds is a detailed picture of what students actually do when they attempt that transition, and how consistently they fall short of it.

Howard, Serviss and Rodrigue (2010) examined the researched writing of eighteen college students and found that none of the papers included a summary of any source. Instead of writing from sources, students were writing from individual sentences selected from those sources. They located a sentence, reproduced or lightly altered it, and moved on. The central question their study raises is one that goes unanswered in the writing itself: does the student understand what has been read? The behaviour they documented, working at the level of the sentence rather than the level of meaning, places the writer in a position where genuine engagement with a source idea never has to occur.

Keck (2014) provides empirical detail of the same pattern across a larger corpus of student summaries. Novice writers, both first-language and second-language, relied more heavily on reproduced source language than experienced writers. Their paraphrases remained close to the original in wording and structure, representing what Keck describes as near copies rather than substantive revisions. The progression from near copy to substantial revision is not merely a linguistic one. It reflects the degree to which the writer has processed the source idea rather than borrowed its surface form.

Shi (2021) examined the same territory from a different angle, focusing on what professors considered acceptable paraphrasing. The professors in her study consistently required evidence that the student had understood and interpreted the source, not merely reproduced it with surface changes. What they were looking for, though they did not use this term, was evidence of content transformation: proof that the writer had engaged with the idea rather than transferred the words. Pecorari (2003) showed that much inappropriate source use stems not from an intention to deceive but from an absence of the means to do otherwise. Students who patchwrite, reproducing source language with minor alterations, are not always avoiding the work of understanding.

Across these bodies of research, a consistent finding emerges. Students who work too closely with source language are not simply making a technical error in citation or paraphrase. In many cases, the issue is developmental rather than deceptive. When reading is reduced to answer-hunting, students learn to treat the text as a container of ready-made responses rather than as material for thought (Gallagher, 2009). What their writing reveals is the absence of an intervening process: the one by which a retrieved idea is held, examined, and rebuilt as the writer's own understanding before it becomes the writer's own words. That process has not been sufficiently examined as a unified educational process in its own right.

2.4 From Retrieval to Understanding

The three bodies of literature reviewed so far establish that reading comprehension requires active construction of meaning, that reading and writing meet in source-based tasks as a single meaning-making process, and that students working from source texts frequently operate at the level of retrieval rather than transformation. What remains to be established is why retrieval is not enough, and what the literature on writing and understanding reveals about the demand that source-based tasks actually make of a learner. This is where the argument arrives at its theoretical core.

Bereiter and Scardamalia (2013) distinguished between two fundamentally different approaches to written composition. In knowledge telling, the writer moves directly from retrieved content to written output. The topic and genre provide the retrieval cues; what comes to mind is written down; the text grows through a process of successive addition rather than problem-solving. The approach is efficient and produces text that can appear complete and structurally correct. What it does not produce is writing that has been genuinely worked through. In knowledge transforming, by contrast, the writer operates across two spaces simultaneously: a content space, in which beliefs and understanding are examined, and a rhetorical space, in which the demands of the written text are negotiated. These two spaces interact. A decision about how to express an idea raises a question about what the idea actually means; that question, once answered, changes what is written. It is through this interaction that writing becomes a process of building understanding rather than a process of depositing what is already known.

The student observed in the 2026 classroom was operating in the manner that Bereiter and Scardamalia (2013) describe as knowledge telling. The source text provided the retrieval cue. The scaffold provided the structural prompt. The sentence moved from page to page without passing through the writer's own understanding. The content space was never entered.

Anderson and Krathwohl (2001) provide a complementary account from the perspective of educational taxonomy. In their revision of Bloom's original framework, remembering and understanding are distinct operations. Remembering involves retrieving relevant knowledge from long-term memory. Understanding involves constructing meaning from what has been retrieved, making sense of it, being able to explain or interpret it in one's own terms. A learner who retrieves correctly but cannot yet explain, interpret, or apply what has been retrieved has completed the first operation but not the second. The source-based examination task requires both. The assessment data from the 2026 cohort's first internal assessment, a school-based comprehension paper, showed students performing the first with increasing reliability whilst the second remained inconsistent.

Wiggins and McTighe (2005) make this visible from a curriculum and assessment perspective. Genuine understanding, they argue, is demonstrated when a learner can explain a concept in their own words, interpret its significance, and apply it in a new context. These are not simply higher demands placed on the same knowledge. They are qualitatively different operations that require the learner to have done something with the retrieved information: processed it, inhabited it, made it theirs.

What Bereiter and Scardamalia (2013), Anderson and Krathwohl (2001), and Wiggins and McTighe (2005) each describe, from different disciplinary positions and with different frameworks, is the same essential gap. This gap requires the learner to convert what has been retrieved into meaning that is genuinely held before that meaning can be expressed in written language.

Between the moment a student locates an idea in a text and the moment that idea appears as expressed meaning in a written response, something must happen. That something is what this paper identifies as idea transformation: the process by which retrieved information is converted into understood meaning before it is committed to writing. The discussion that follows returns to the classroom evidence, drawn from both the cohort's first internal comprehension assessment and a subsequent directed writing lesson, to examine what idea transformation looks like in its absence, and what its absence means for teaching.

3. Discussion

This paper is a conceptual practitioner inquiry. The classroom evidence presented here is not formal research data gathered through a planned methodology. The evidence is used in a different and specific way: to make visible a classroom process that existing literature has approached from several directions but has not named as a unified teaching concern. The twenty-nine scripts from the cohort's first internal comprehension assessment are drawn on as illustrative evidence of a pattern that the theoretical framework helps explain. The individual student moment from the directed writing lesson is used in the same way: not as a case study in the methodological sense, but as a precise and concrete instance that makes the argument visible at the level of a single learner at a single desk. The paper's contribution is therefore conceptual. It proposes a name and a theoretical position for a classroom process that teachers observe regularly but have not had precise language to address directly.

3.1 *What the Evidence Shows*

The evidence presented in this section moves across three levels, each doing different work: the examination demand, the cohort pattern, and the individual classroom moment. The Cambridge O Level English Language Principal Examiner Report (Cambridge Assessment International Education, 2025) establishes the external expectation that candidates should use source material rather than reproduce it. The cohort evidence shows that this expectation was not yet being met consistently by the 2026 O Level Intensive Programme students. The individual student example then makes the missing process visible in a single, concrete moment.

The report does not merely identify a general weakness in student writing. It names the demand of the task. Directed writing requires candidates to read one or more source texts and produce a piece of writing, such as a letter, report, or speech, that draws on and develops the information provided. Success therefore depends not only on locating the correct information, but on reshaping that information for a new purpose, audience, and form. In comprehension and summary tasks, weaker responses were characterised by lifting from the source and narrative retelling rather than explanation and development. The report provides the examination standard against which the classroom evidence in this paper can be understood: the examination expects students to do something with what they find, not simply place it on the page.

The cohort evidence shows how this demand appeared in practice. Across the twenty-nine scripts of the 2026 O Level Intensive Programme first assessment, a school-based comprehension assessment administered at a defined point in a targeted intervention programme for students working towards examination resit, the same pattern appeared repeatedly. Students were able to locate relevant information in the source texts. They could identify useful points and position them within a written structure. Yet their responses frequently remained close to the wording of the source. Where explanation was attempted, it often restated rather than developed the original point. Where independent expression was expected, the language of the source text reappeared with minimal change. Sentences from the source texts appeared in student responses with small grammatical adjustments but without any visible attempt to restate the idea or extend it. The difficulty was not that students could not find information. The difficulty was that retrieved information did not consistently become understood meaning before it entered the written response. This is what the first assessment, as a school-based assessment, made visible: not an absence of effort, but an absence of a specific process.

This pattern became visible with particular clarity in the work of one student during a subsequent directed writing lesson. The student read Text A, identified a relevant sentence, and highlighted it. She then turned to the paragraph scaffold and copied the highlighted sentence into her response with very little change in wording. The scaffold provided a sentence starter designed to guide the student into expressing an idea in a structured paragraph. At this point, the student had already located the relevant idea. The next expected move was not writing, but transformation: holding the idea, making sense of it, and preparing to express it in her own language. Instead, she moved directly from the highlighted sentence to the scaffolded paragraph. Looking across the classroom at that moment, the same pattern was visible in other responses: highlighted sentences transferred into scaffolded paragraphs with little evidence of intervening thought.

The significance of this moment lies in what did not happen. The student had not failed to read the text. She had not failed to identify relevant material. She had not failed to place the point in the correct part of the response. What was missing was the intermediate process between retrieval and expression. The idea moved from the source text to the answer sheet without passing visibly through the student's own understanding.

The evidence points towards a problem this paper identifies as distinct and overlooked. Copying, weak paraphrasing, and limited sentence control are the visible symptoms. The deeper issue is the absence of a middle step in source-based writing: the process this paper has named idea transformation.

3.2 Reading the Evidence through the Theoretical Framework

Locating source material and placing it into a written response is not the same as transforming it. The theoretical framework helps explain why that distinction matters. The issue is not simply that students copy. The more important concern is that the task sequence may allow students to stop at retrieval, whilst the examination actually requires them to move beyond retrieval into understood expression.

Raymond's (1988) discussion of cloze procedure is useful here because it shows how a reading task can appear simple whilst concealing more complex demands. In a cloze task, the student reads, identifies a missing word, supplies an answer, and moves on. The task may support scanning, prediction and reconstruction, but it can also encourage a narrow form of completion if it is used mechanically. The risk she identifies is that a task may appear to develop comprehension whilst stopping short of the meaning-making that comprehension requires. In the same way, a directed writing scaffold may help students locate and place information, yet still leave unaddressed the process by which an idea is held, tested, understood and expressed.

Kintsch (1988) helps explain why this matters. Reading comprehension is not simply a matter of recognising words or locating information. It requires the construction of a mental representation in which ideas are connected, interpreted and understood in relation to one another. When a student moves directly from a highlighted sentence to a written response, what is missing is not the words. What is missing is the constructed meaning that should stand between reading and writing. Anderson and Krathwohl (2001) sharpen this distinction further. In their revision of Bloom's taxonomy, remembering and understanding are distinct processes. Retrieving information from a source text is an act of remembering. Converting that information into an idea the learner can explain, develop and express is an act of understanding. The student in Section 3.1 completed the first process. The second process did not visibly occur.

Bereiter and Scardamalia (2013) name the writing behaviour that results. Knowledge telling describes the practice of retrieving what is known or found and committing it directly to the page without the transformation that genuine composition requires. The student who highlighted a sentence and copied it into a scaffolded paragraph was not yet composing in the fuller sense. She was telling. This behaviour is consistent with what Howard, Serviss and Rodrigue (2010) describe as patchwriting: the reproduction of source language with only minimal alteration, a pattern more accurately understood as a developmental limitation than as deliberate copying. The idea moved from the source to the answer sheet without passing through the generative process that writing, at its most purposeful, demands.

Rahman et al. (2021) extend this concern from task design to assessment culture, drawing on the established concept of washback: the effect that high-stakes assessments have on teaching and learning (Alderson & Wall, 1993). Their work on washback shows that assessment systems can shape what teachers emphasise and what students learn to value. When assessment pressure rewards visible completion, correct selection and examination performance, classroom practice can become increasingly task-driven. Students may learn to look for the answer, reproduce the answer and move on, because those behaviours appear to satisfy the immediate demand. In such a context, the absence of idea transformation is not only an individual weakness. It becomes part of a wider pattern in which assessment conditions may encourage students to complete the task without fully engaging with the meaning of the source.

Gabut et al. (2026) deepen the argument by distinguishing surface learning from deeper engagement. Surface learning relies on memorisation, reproduction and minimum task completion. Deep learning requires students to connect, organise and understand ideas. The student who copies a highlighted sentence into a scaffolded paragraph is not necessarily unwilling to think. She may be using the strategy that the learning environment has made most familiar: find the relevant information, place it correctly, and produce an acceptable-looking response. From the outside, the answer appears organised. From the inside, the idea has not yet been transformed.

These sources help interpret the evidence in a way the evidence alone cannot. Raymond (1988) shows what a task can miss. Kintsch (1988) and Anderson and Krathwohl (2001) explain what is missing at the level of meaning and understanding. Bereiter and Scardamalia (2013) name the resulting writing behaviour. Rahman et al. (2021) show how assessment pressure can reinforce incomplete task behaviours. Gabut et al. (2026) show how students may come to rely on surface strategies when those strategies appear to work. The classroom moment described above therefore points to more than copying. It reveals an unnamed gap in the teaching sequence: students are taught to find information and to write answers, but they are not always taught how to convert the information they have found into meaning they genuinely understand.

This is why idea transformation matters as a concept. If the missing process is not named, it is difficult to teach directly. Students may continue to be corrected for copying without being shown what should happen instead. The problem is therefore not only that students reproduce source language. The deeper problem is that the intermediate act between retrieval and written expression has remained insufficiently visible as a teaching objective.

3.3 What the Evidence Means for Teaching

The evidence and the theoretical framework have established what is missing. What remains is the practical question: if students are stopping at retrieval, what must teaching insert between finding and writing?

The answer is not simply a new technique. It is a different understanding of what the writing task requires. Elbow (1973) argues that writing is not the transcription of ideas that already exist in finished form. It is the process through which meaning grows. When a student moves directly from a highlighted sentence to a written response, she is treating writing as transcription. The idea is already there, in the source text. Her task, as she understands it, is to place it correctly. What Elbow's argument makes visible is that this understanding of writing misses the generative middle: the space in which the writer holds an idea, works with it, and arrives at an expression that is genuinely her own. That space is not incidental to purposeful writing. It is where transformation happens.

The pedagogical implication is therefore specific. Teaching must make the generative middle visible and necessary. One way to do this is to remove the source sentence from view at the point of writing. If the student cannot see the original, she must reconstruct the idea in her own language. Reconstruction is not paraphrase practised as a technique. It is the act of holding meaning long enough to express it. That is where idea transformation becomes observable and, importantly, where it becomes teachable as a classroom step rather than an assumed outcome.

Khan et al. (2025) caution, however, that task redesign alone is insufficient. Their study of assessment reform in Malaysia shows that even well-designed tasks intended to promote deeper engagement with meaning can be neutralised by the surrounding system. When assessment conditions continue to reward completion and correct selection, when teacher training does not address the gap between retrieval and expression, and when institutional expectations remain focused on examination performance, students and teachers alike will orient their practice towards the visible surface requirement. The teachable middle step that this paper identifies requires not only different tasks but different teacher understanding of what those tasks are for, and assessment conditions that make meaningful engagement with the idea visible and valued.

4. Conclusion

This paper has argued from a specific and bounded position. The evidence is drawn from one cohort, in one institution, at one point in the programme. The framework is built from existing literature rather than from a designed empirical study. These are not weaknesses to apologise for. They are the honest boundaries of a conceptual paper that has identified a focused and specific problem: the absence, in classroom practice and in the existing literature, of a named process connecting reading to writing. What the paper opens is a question for further investigation, specifically, how idea transformation can be taught, observed, and strengthened across different source-based writing contexts within English language education. That investigation is the work that follows.

What the paper has established, within those boundaries, is this. Between the moment a student locates an idea in a source text and the moment that idea appears in a written response, something must happen. That something has a name. Idea transformation is the process by which a learner converts retrieved information into understood meaning before committing that meaning to writing. It is not a new mechanism of the mind. It is an overlooked educational process, observable in classroom practice and supported by existing theory, that sits between retrieval and expression and has not previously been examined as a unified teaching concern.

The first research question asked what idea transformation is and where it sits in the sequence connecting reading to writing. The answer the paper has built is as follows. Reading comprehension, as Kintsch (1988) established, is not completed by locating a relevant sentence. It requires the construction of meaning from what has been read. Writing from sources, as Spivey (1990) and the reading-writing relationship literature demonstrate, is not a transfer of that located material onto the page. It is a continuation of the same meaning-making process under new demands. Idea transformation is what must occur between these two acts: the moment at which the learner stops receiving an idea and begins doing something with it. In Anderson and Krathwohl's (2001) terms, it is the passage from remembering to understanding. In Bereiter and Scardamalia's (2013) terms, it is the entry into the content space that knowledge transforming requires and knowledge telling bypasses. It sits between retrieval and expression, and it is precisely there that the students in the 2026 cohort were stopping short.

The second research question asked what the absence of idea transformation looks like in the classroom, and what this suggests for teaching. The answer is visible in two kinds of evidence. Across twenty-nine scripts in the cohort's first internal comprehension assessment, retrieved information entered written responses with minimal processing. The language of the source text reappeared with small adjustments but without any visible attempt to restate, develop, or extend the idea. In a subsequent directed writing lesson, a single student made the missing process concrete: she read, highlighted, and copied, moving directly from the source text to the scaffolded paragraph without passing through understanding. What was absent was not effort and not attention. What was absent was the intermediate act between finding and writing.

For teaching, this absence points to a specific and addressable gap. Elbow (1973) establishes that writing is not transcription but generation: meaning grows through the writing process rather than preceding it. If that is so, then teaching must make the generative middle visible and necessary. The practical implication is not a new technique but a reorientation of what the writing task is understood to require. Removing the source sentence from view at the point of writing is one way to make the intermediate act unavoidable: the student must reconstruct the idea in her own language, which is precisely where idea transformation becomes observable and teachable. Khan et al. (2025) caution, rightly, that task redesign alone is insufficient where assessment conditions continue to reward visible completion over meaningful engagement. The teaching implication this paper identifies therefore requires not only different tasks but different understanding of what those tasks are for.

The contribution this paper makes is not empirical. It is conceptual. It proposes a name for a process that classroom teachers observe repeatedly but have not had precise language to describe or address directly. Students are corrected for copying. They are told to use their own words. What they are rarely shown is what must happen in the space between finding an idea and writing it down. Naming that space, and positioning it within an established theoretical framework, is the first step towards teaching it deliberately. That is what this paper has attempted to do.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest: The author declares no conflicts of interest with respect to the research, authorship, and/or publication of this article.

Informed Consent Statement/Ethics Approval: Not applicable.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129. <https://doi.org/10.1093/applin/14.2.115>
- 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.
- Bereiter, C., & Scardamalia, M. (2013). *The psychology of written composition*. Routledge. <https://doi.org/10.4324/9780203812310>
- Cambridge Assessment International Education. (2025). *Cambridge O Level English Language 1120 principal examiner report for teachers*. Cambridge Assessment International Education.
- Elbow, P. (1973). *Writing without teachers*. Oxford University Press.
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. <https://doi.org/10.58680/coc198115885>
- Gallagher, K. (2009). *Readicide: How schools are killing reading and what you can do about it*. Stenhouse Publishers.
- Gabut, D. M. P., Rojo, M. E. M., Casas, W. C., Tondo, R. C., Jr., Mahilum, R. M., & Mutya, R. C. (2026). Comparing surface and deep learning strategies and their relationship to high school students' academic performance in science. *European Journal of Educational Research*, 15(1), 199–210. <https://doi.org/10.12973/eu-jer.15.1.199>
- Howard, R. M., Serviss, T., & Rodrigue, T. K. (2010). Writing from sources, writing from sentences. *Writing and Pedagogy*, 2(2), 177–192. <https://doi.org/10.1558/wap.v2i2.177>
- Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, 25, 4–22. <https://doi.org/10.1016/j.jslw.2014.05.005>
- Khan, A., Hassan, N., & Cheng, L. (2025). Investigating the contextual factors mediating washback effects of a learning-oriented English language assessment in Malaysia. *Language Testing in Asia*, 15, Article 20. <https://doi.org/10.1186/s40468-025-00359-8>
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163–182. <https://doi.org/10.1037/0033-295X.95.2.163>
- Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*, 12(4), 317–345. <https://doi.org/10.1016/j.jslw.2003.08.004>
- Rahman, K. A., Seraj, P. M. I., Hasan, M. K., Namaziandost, E., & Tilwani, S. A. (2021). Washback of assessment on English teaching-learning practice at secondary schools. *Language Testing in Asia*, 11, Article 12. <https://doi.org/10.1186/s40468-021-00129-2>
- Raymond, P. M. (1988). Cloze procedure in the teaching of reading. *TESL Canada Journal*, 6(1), 91–97. <https://doi.org/10.18806/tesl.v6i1.544>
- Shanahan, T. (1988). The reading-writing relationship: Seven instructional principles. *The Reading Teacher*, 41(7), 636–647. <https://doi.org/10.2307/20199888>
- Shi, L. (2021). Professors' views of content transformation in students' paraphrasing. *Journal of Response to Writing*, 7(2), 112–144. <https://scholarsarchive.byu.edu/journalrw/vol7/iss2/5>
- Spivey, N. N. (1990). Transforming texts: Constructive processes in reading and writing. *Written Communication*, 7(2), 256–287. <https://doi.org/10.1177/0741088390007002004>
- Tierney, R. J., & Pearson, P. D. (1983). Towards a composing model of reading. *Language Arts*, 60(5), 568–580. <https://doi.org/10.58680/la198326307>

- Tovani, C. (2000). *I read it, but I don't get it: Comprehension strategies for adolescent readers*. Stenhouse Publishers.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). ASCD.
- Wittrock, M. C. (2010). Learning as a generative process. *Educational Psychologist*, 45(1), 40–45.
<https://doi.org/10.1080/00461520903433554>
Original work published 1974

Perceptions And Acceptance of Artificial Intelligence in Science Education Programmes: Voices of Pre-Service Science Teachers

Kwaku Darko Amponsah¹, Joseph Adu-Boahen², Priscilla Commey-Mintah³, Eliot Kosi Kumassah⁴,
Raphael Forster Ayithey⁵, John Nketsiah⁶

¹ Department of Teacher Education, School of Education and Leadership, University of Ghana,
Legon, Accra, Ghana. ORCID: 0000-0002-7824-6516

Email: kdamponsah@ug.edu.gh

² Department of Teacher Education, School of Education and Leadership, University of Ghana, Legon, Accra,
Ghana. ORCID: 0009-0008-6236-0366

Email: adujoseph053@gmail.com

³ Department of Teacher Education, School of Education and Leadership, University of Ghana, Legon, Accra,
Ghana. ORCID: 0000-0002-6575-4335

Email: priscom2000@gmail.com

⁴ Department of Teacher Education, School of Education and Leadership, University of Ghana, Legon, Accra,
Ghana. ORCID: 0000-0001-6203-4992

Email: ekummassah@ug.edu.gh

⁵ Department of Teacher Education, School of Education and Leadership, University of Ghana, Legon, Accra,
Ghana. ORCID: 0009-0007-1170-4686

Email: rayithey@ug.edu.gh

⁶ Department of Science, Accra College of Education, P. O. Box 221, Legon- Accra, Ghana.

ORCID: 0000-0002-5811-5463

Email: j.nketsiah@acce.edu.gh

Correspondence: Priscilla Commey-Mintah, Department of Teacher Education, School of Education and
Leadership, University of Ghana, Legon, Accra, Ghana. Email: priscom2000@gmail.com

Abstract

This study examined pre-service science teachers' perceptions, acceptance, and use of artificial intelligence (AI) in science education programmes in Ghana. Guided by the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Theory of Planned Behaviour (TPB), the study explored the AI tools frequently used by pre-service science teachers, the purposes for which these tools are employed, and their perceptions, behavioural intentions, and actual use of AI in science teaching and learning. A quantitative cross-sectional survey design was adopted, involving 380 pre-service science teachers selected from science education programmes. Data were collected using a structured questionnaire and analysed using descriptive and inferential statistics. The findings revealed generally positive perceptions of AI, with participants recognising its potential to improve teaching quality, instructional effectiveness, and student learning outcomes. ChatGPT emerged as the most frequently used AI tool, while research, content explanation, and lesson planning were the primary purposes for its use. Although

participants demonstrated positive attitudes, favourable effort expectancy, and strong behavioural intentions toward AI adoption, actual classroom use remained moderate. The study concludes that effective AI integration requires enhanced digital infrastructure, targeted professional development, and supportive institutional policies to bridge the gap between intention and practice.

Keywords: Artificial Intelligence, Science Education, Pre-Service Science Teachers, Technology Acceptance, Behavioural Intention, AI integration, Ghana

1. Introduction

The rapid expansion of artificial intelligence (AI) is profoundly transforming education, offering significant opportunities and challenges for teaching and learning. AI systems, characterised by their capacity to learn, adapt, synthesise information, and self-correct, have become central to educational discussions. This transformation is especially evident in science education, where AI technologies—including generative models like ChatGPT, adaptive tutoring systems, and automated assessment tools—are poised to revolutionise the delivery of scientific concepts, facilitate inquiry-based learning, and increase student engagement with complex subject matter (Mustafa, 2024; T Kotsis, 2025). Previous research in science education has demonstrated that technology-supported and collaborative instructional approaches can significantly improve students' conceptual understanding of difficult scientific concepts, highlighting the potential of innovative technologies to enhance science learning outcomes (Amponsah & Ochonogor, 2018). Educators' perceptions—encompassing cognitive and emotional evaluations of AI's benefits, risks, and suitability—along with their acceptance, reflected in intentions and actual usage, critically influence the meaningful integration of AI in classrooms. This observation is consistent with findings from Ghanaian secondary schools, where teachers' perceptions of technology initiatives significantly influenced their instructional practices and willingness to integrate digital tools into teaching and learning processes (Asoma et al., 2024). The issue is pressing, as AI advancements outpace pedagogical adaptations and regulatory measures, potentially exacerbating disparities in preparedness and equity (Cabero-Almenara et al., 2024; J. Kim, 2025).

Current research consistently identifies performance expectancy and effort expectancy as key determinants of teachers' intentions to adopt AI. Investigations involving science and STEM educators reveal recognition of AI's capacity to enhance the delivery of complex content and alleviate routine tasks. Factors such as AI familiarity, teaching experience, academic self-efficacy, and instructional beliefs further shape readiness for AI integration (Roshan et al., 2024; Sova et al., 2024). However, the literature reveals tensions: while quantitative surveys often show generally favourable attitudes toward AI, qualitative and mixed-methods studies emphasise concerns about data privacy, algorithmic transparency, the potential dehumanisation of education, and fears that AI will replace teachers (Alghamdy, 2023; Dignum, 2023). Furthermore, scholars argue that traditional acceptance frameworks may inadequately address the ethical and epistemological complexities unique to AI. A significant research gap persists: most studies focus on general educator populations in Western and Asian contexts, with limited quantitative inquiry into pre-service science teachers' perspectives in Ghana (Akanzire et al., 2025; Eusebio et al., 2025) (Ishmuradova et al., 2025; Awofala et al., 2025; Adelana et al., 2024).

This study addresses this void through a quantitative analysis of pre-service science teachers' perceptions and acceptance of AI in Ghana. Guided by the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Theory of Planned Behaviour (TPB), this study examined performance expectancy, attitude, perceived behavioural control, behavioural intention, and actual use. This research explores the scope and patterns of AI engagement among this population.

1.1. Statement of the Problem

Although AI tools such as ChatGPT and Microsoft Copilot are becoming increasingly accessible, little is known about how pre-service science teachers in Ghana perceive their pedagogical value, ethical implications, and applicability for lesson planning, instruction, assessment, and student engagement. Furthermore, much of the existing research on AI adoption in education has focused on general teaching contexts, with inadequate emphasis

on science education, where inquiry-based learning and conceptual mastery are fundamental (Agormedah et al., 2022; Agyei & Voogt, 2010; Akyeampong, 2017). This gap is compounded by contextual challenges such as infrastructural limitations, limited professional development opportunities, and concerns related to academic integrity and ethical AI use. Similar challenges have been reported in studies examining technology integration in Ghanaian schools, where teacher readiness and perceptions were found to influence the successful implementation of educational technologies (Asoma et al., 2024). As a result, these may shape teachers' readiness and acceptance of AI. Therefore, there is a clear need for an empirical investigation into pre-service science teachers' perceptions and acceptance of AI in science education programmes in Ghana to inform teacher preparation, policy, and pedagogical practice. In line with the above, this study aims to:

- i. Identify AI tools frequently used by pre-service science teachers in Ghana for teaching and lesson preparation.
- ii. Explore the primary purposes for which pre-service science teachers use AI in science education programmes.
- iii. Assess pre-service science teachers' perceptions, intentions, and actual use of AI in science education.

1.2. Research Questions

1. Which AI tools are frequently used by pre-service science teachers in Ghana?
2. For what purposes do pre-service science teachers use AI in science education programmes?
3. What are pre-service science teachers' perceptions, behavioural intentions, and actual use of AI in science education?

2. Literature Review

2.1. Theoretical Framework

Implementing artificial intelligence (AI) in science teaching is both a technological advancement and a cognitive activity that involves teachers' beliefs, judgments, and sense of agency. While most studies have focused on whether teachers view AI as beneficial and whether it is uncomplicated to integrate, fewer have examined the cognitive processes that inform such teaching behaviours. To bridge the gap between AI teaching perception and practice, this study adapts the Unified Theory of Acceptance and Use of Technology (UTAUT) by Venkatesh et al. (2003) and the Theory of Planned Behaviour (TPB) by Ajzen (1991), which has been expanded in subsequent work. The UTAUT framework offers a basic understanding of how people assess new technologies. Unlike previous acceptance models, such as the technology acceptance model (Davis, 1989), UTAUT draws on all eight major theories to explain technology acceptance and use (Venkatesh et al., 2003). In this model, performance expectancy reflects teachers' views of AI's ability to improve their effectiveness and students' thinking skills in learning science. Evidence from science education research suggests that instructional innovations that enhance collaboration and active engagement can significantly improve students' conceptual understanding and learning outcomes, thereby strengthening teachers' perceptions of their usefulness (Amponsah & Ochonogor, 2018). Recent studies in educational technology and AI have consistently shown that performance expectancy is a major factor influencing teachers' willingness to work with AI (Kim et al., 2024; Teo & Noyes, 2012; Tram, 2024). Teachers' willingness to use AI technologies has also been influenced by their perceived effort expectancy, with AI and technical training falling short of expectations (Fang et al., 2025). Recent improvements to UTAUT have highlighted the importance of the above factors in assessing teachers' responses to advanced technologies with data (AI) (Venkatesh & Thong, 2016). By contrast, many researchers believe that UTAUT is good at predicting a person's intention but weak at explaining the behaviour that the intention is materialised. This limitation has prompted the need for integrative frameworks that focus on behavioural mechanisms rather than direct-effect explanations.

The theory of planned behaviour (TPB), introduced by Ajzen (1991) and later expanded (Ajzen & Schmidt, 2020), attempts to address this by putting a model to the thought processes that lead to that behaviour. Key to the TPB is the claim that a cognitive commitment to the action is the strongest predictor of its occurrence. In educational

technology, the TPB has been cited to explain the divergence among teachers in their adoption of digital and AI-driven technologies (Bosnjak et al., 2020). With the TPB, attitudes toward AI refer to teachers' positive or negative evaluations of using AI in science teaching, including its perceived value, relevance, and instructional efficiency. In the TPB, PBC is a construct introduced by Ajzen (1991) to explain the influence of non-volitional factors on behaviour, and it refers to the extent to which a teacher believes they can utilise AI to teach effectively, given the available resources, skills, and supportive structures within the institution. Recent studies on TPB and technology adoption have focused on the other impact of behavioural intention and have found that attitudes and PBC do not directly impact technology use but influence other factors that determine it (Naskar & Lindahl, 2025). This applies to the context of science education, where teachers may have a positive conceptual attitude toward AI but lack the external factors to translate that attitude into practice.

2.2 Conceptual Framework

Behavioural intention (BI) refers to the conscious effort one exerts to perform a specific behaviour and is one of the strongest predictors of future occurrences of similar behaviours in the fields of health and technology (Ajzen, 1991; Venkatesh et al., 2003). Between the Theory of Planned Behaviour and the Unified Theory of Acceptance and Use of Technology, behavioural intention is placed closest to the behaviours in question, and the influences of attitudes, norms, perceived behavioural control, and the performance and effort expected are channelled primarily through behavioural intention. Venkatesh et al. (2003) built upon previous research to separate the concepts of BI and behavioural expectation through the lens of social influence and facilitating conditions, which they posit are more essential to expectations in technology contexts. There is considerable evidence supporting the intention–behavior link. Budu et al. (2018) demonstrated this phenomenon in Ghanaian tertiary institutions, where BI was identified as the mediating factor in the relationship between self-efficacy and the actual use of e-learning.

Performance expectancy (PE) reflects the extent to which someone views a given system as a contributor to improved job performance. It utilises foundational constructs, such as perceived usefulness from the technology acceptance model, extrinsic motivation, and relative advantage from innovation diffusion theory (Venkatesh et al., 2003). It is the primary driver of perceived efficiency and effectiveness of technology and is widely cited as the most influential factor in behavioural intention. Recent empirical studies have continued to extend this role across various contexts. Most notably, PE was the strongest predictor (among all constructs studied) of mobile learning among postgraduate students in Nigeria (Oyewole, 2018).

Effort expectancy (EE) denotes the perceived ease of utilising a given system. It indicates the levels of complexity and usability hypothesised in the technology acceptance model (TAM) and in the UTAUT (Venkatesh et al., 2003). Systems perceived as complex and difficult to use may impose cognitive and physical costs that discourage use, even when the value is clear; in contrast, a system considered easy to use may reduce the psychological costs associated with it and increase the intention to adopt it. While there is a general context that supports the relevance of EE, there are outliers within it. Oyewole (2018) acknowledged a notable positive correlation between EE and smartphone use, specifically in mobile learning ($r = 0.724$), highlighting the value of ease of use in a population with a low level of digital literacy. In contrast, Bayaga and du Plessis (2024) found no significant positive correlation between EE and behavioural intention among university students in South Africa, suggesting that control may be moderated by the context's hierarchy, structure, or culture.

Perceived behavioural control (PBC) is defined as the ease with which an individual believes they can perform a behaviour. PBC captures both internal barriers and facilitators (e.g., skills and abilities) and external barriers and facilitators (e.g., resources, opportunities, and organisational backing) (Ajzen, 1991). PBC is conceptually similar to Bandura's notion of self-efficacy, except that it considers external environmental barriers. There is an extensive body of literature that illustrates the predictive power of PBC. Research involving Ghanaian pre-service teachers has also shown that individual cognitive attributes and learner characteristics influence confidence, competence, and educational performance, factors that are conceptually linked to perceptions of behavioural control and technology adoption readiness (Budu et al., 2022).

Attitude (AT) indicates a person's general positive or negative assessment of a behaviour; in this case, how positive or negative a user evaluates a system. This evaluation emerges from a system user's cognitive assessments of potential consequences and emotional responses (Ajzen, 1991). Such evaluations relate to salient behavioural beliefs. Positive attitudes toward educational innovations have similarly been observed among Ghanaian teachers participating in game-based learning initiatives, with favourable perceptions contributing to a greater willingness to adopt learner-centred pedagogical approaches (Yeboah et al., 2025). These beliefs are outcome expectations, such as time savings and improvements in the quality or efficacy of technological advancements, and are positively associated with the perceived importance of these improvements. According to Budu and Mireku (2018), participants' positive attitudes influenced their willingness to engage the public.

3. Methodology

3.1 Research Design and Selection of Participants

This study utilised a quantitative cross-sectional research design. This design allows for a detailed description of pre-service science teachers' perceptions, acceptance, and use of artificial intelligence in science education. The design allows for the collection of data at one point in time and provides a strong framework for comparisons between groups (Creswell & Creswell, 2023).

3.2 Population and Participant Selection

A purposive sampling technique was employed to select 380 pre-service science teachers out of a total population of 450 who were actively engaged in science education programmes. This sampling approach ensured that participants were well-positioned to provide informed insights into the integration of artificial intelligence (AI) in science education programmes, thereby capturing a broad range of experiences and perceptions (Creswell & Creswell, 2023).

3.3 Data Collection and Instrument

Data were collected using a self-developed, structured questionnaire comprising seven sections designed to examine pre-service science teachers' perceptions, acceptance, and use of artificial intelligence (AI) in science education. The first section captured respondents' demographic characteristics. Sections 2 to 6 measured key AI adoption constructs, including performance expectancy (e.g., using AI will improve the quality of my science lessons), effort expectancy (e.g., AI applications are easy to learn and use), attitude towards AI (e.g., I feel positive about using AI in my science lessons), perceived behavioral control (e.g., confidence in using AI for science teaching), and behavioral intention (e.g., intention to use AI in future science lessons). The final section assessed actual use of AI, focusing on respondents' self-reported use of AI tools in science lessons during the past semester. All construct-related items were measured using a five-point Likert-type scale to capture respondents' levels of agreement.

3.4 Data Analysis Method

Data were analysed using both descriptive and inferential statistical techniques. Descriptive statistics (means and standard deviations) and inferential statistics (one-sample t-tests and associated p-values) were used to summarise teachers' AI usage patterns, perceptions, and behavioural intentions. To test whether the mean responses differed significantly from the neutral midpoint, one-sample t-tests were conducted. Additionally, the measurement model was assessed through reliability and validity analyses, including Cronbach's alpha, composite reliability, and average variance extracted (AVE), to confirm the internal consistency and convergent validity of the constructs. All analyses were performed using the Statistical Package for the Social Sciences (SPSS).

3.5 Ethical Consideration

Informed consent was obtained from all participants, who were made aware of the study's purpose, their right to withdraw, and how their data would be utilised. Confidentiality was maintained by anonymising participant information and securely storing data.

3.6 Validity and Reliability

Table 1: Reliability and Validity of Study Constructs

	Cronbach's Alpha	Composite Reliability (rho_C)	Composite Reliability (rho_A)	AVE
Actual Use of AI	.876	.909	.885	.668
Attitude Towards AI	.934	.948	.936	.752
Behavioural Intentions	.918	.936	.923	.709
Effort Expectancy	.854	.892	.870	.581
Perceived Behavioural Control	.926	.942	.929	.730
Performance Expectancy	.907	.928	.913	.682

Source: Author (2026)

The measurement model demonstrated strong internal consistency and adequate convergent validity. All constructs recorded Cronbach's alpha and composite reliability values above the recommended threshold of 0.70, indicating high reliability. Additionally, the average variance extracted (AVE) values exceeded 0.50 for all constructs, indicating that each construct explained a sufficient proportion of the variance in its indicators. The results confirmed that the measurement model was reliable and valid, providing evidence that the study constructs were suitable for subsequent statistical analyses of pre-service science teachers' perceptions and acceptance of artificial intelligence in science education programmes.

4. Results and Discussion

4.1 Results

Identify AI tools frequently used by pre-service science teachers in Ghana for teaching and lesson preparation.

Which AI tools do you frequently use for teaching or lesson preparation?
140 responses

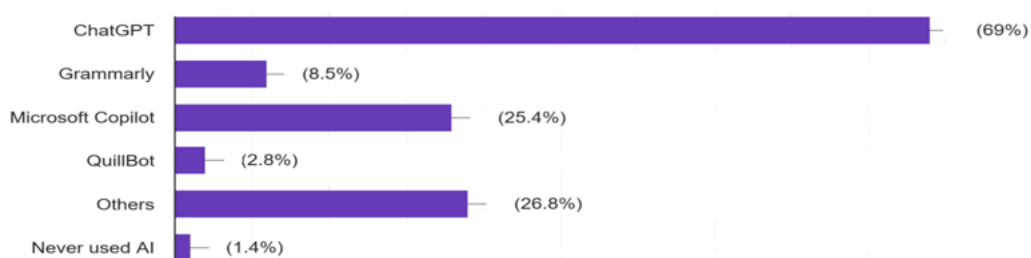


Figure 1: AI Tools Frequently Used by Pre-Service Science Teachers

Source: Author (2026)

The results indicate a high level of AI adoption among pre-service science teachers for teaching and lesson preparation. ChatGPT emerged as the most widely used tool, with 69% of respondents reporting frequent use, underscoring its central role in instructional planning and content development. Microsoft Copilot (25.4%) and other AI tools (26.8%) also recorded moderate usage, suggesting a growing openness to multiple AI platforms. In contrast, Grammarly (8.5%) and QuillBot (2.8%) were used by relatively few teachers, implying that AI applications focused on language editing are less integral to science teaching. Importantly, only 1.4% of respondents reported never having used AI, reflecting widespread exposure to and engagement with AI technologies in science education programs.

Explore the primary purposes for which pre-service science teachers use AI in science education programmes.

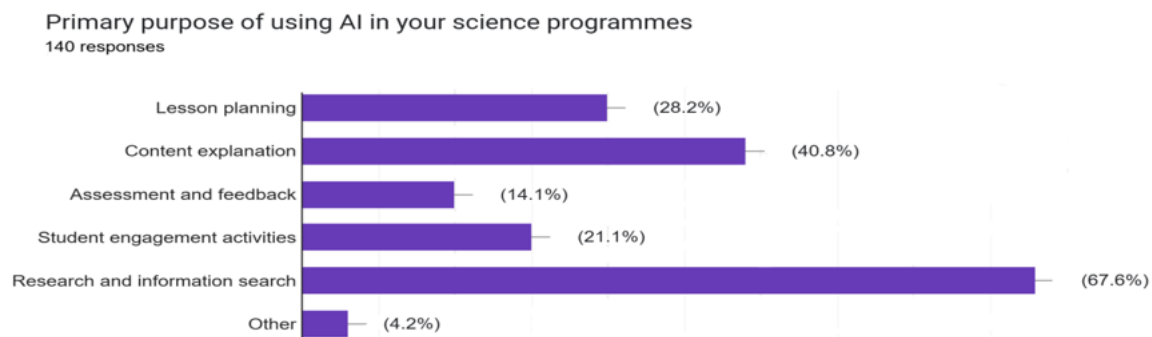


Figure 2: Primary Purpose of Using AI by Pre-Service Science Teachers

Source: Author (2026)

The chart shows that the majority of teachers use AI primarily for research and information searching (67.6%). The second most common use is content explanation (40.8%), followed by lesson planning (28.2%). Student engagement activities accounted for 21.1%, while assessment and feedback were the least selected among the main categories (14.1%). Only 4.2% indicated other purposes for using AI. These results suggest that teachers are mainly using AI as a support tool for research and content delivery, rather than for assessment or student interaction.

Assess pre-service science teachers' perceptions, intentions, and actual use of AI in science education.

Table 2: Descriptive Statistics on Performance Expectancy

	Mean	SD	T-value	P-value
I believe using AI in science teaching will improve the quality	4.04	.933	36.525	< .001
AI tools will make my teaching more efficient	4.04	1.048	32.502	< .001
Using AI in the classroom will help students understand science	3.93	.961	34.455	< .001
AI can help me achieve better learning outcomes for my students	4.01	.886	38.163	< .001
Incorporating AI into teaching will make my work more effective	4.13	.955	36.406	< .001
AI adoption is likely to improve my overall teaching performance	3.80	1.154	27.764	< .001

Source: Author (2026)

The results indicate that teachers generally have positive perceptions of the integration of artificial intelligence (AI) in science teaching. All items received mean scores above the midpoint of 3.0, reflecting favourable attitudes toward AI adoption. Specifically, teachers agreed that AI could improve the quality of science teaching ($M = 4.04$, $SD = .933$) and make teaching more efficient ($M = 4.04$, $SD = 1.048$). They also believed that the use of AI in the classroom would enhance students' understanding of science ($M = 3.93$, $SD = .961$) and improve learning outcomes ($M = 4.01$, $SD = .886$). Additionally, participants agreed that incorporating AI into teaching would make their work more effective ($M = 4.13$, $SD = .955$) and that AI adoption is likely to improve their overall teaching performance ($M = 3.80$, $SD = 1.154$). One-sample t-test results revealed that all these mean scores were significantly higher than the neutral midpoint, as indicated by statistically significant t-values ($p < .001$), suggesting that teachers' positive perceptions of AI are not due to chance but represent a meaningful trend in the sample.

Table 3: Descriptive Statistics on Effort Expectancy

	Mean	SD	T-value	P-value
I find AI applications easy to learn and use in science teaching.	3.89	.994	32.968	< .001
Interaction with AI tools in the classroom is clear and understandable.	3.90	.848	38.777	< .001
Using AI in science lessons requires little effort.	3.52	1.094	27.130	< .001
AI tools can be used without extensive technical knowledge.	3.69	1.141	27.244	< .001
AI software and applications are user-friendly for teaching purposes	3.89	.871	37.609	< .001
Integrating AI into my existing lesson plans is easy for me.	3.80	.965	33.192	< .001

Source: Author (2026)

The results for effort expectancy indicate that teachers generally perceive AI applications as easy to use and integrate into science teaching. All six items had mean scores above the neutral midpoint of 3.0, suggesting that participants found the AI tools relatively user-friendly. Specifically, teachers agreed that AI applications are easy to learn and use ($M = 3.89$, $SD = 0.994$) and that interacting with AI tools in the classroom is clear and understandable ($M = 3.90$, $SD = 0.848$). They also reported that AI software and applications are user-friendly for teaching purposes ($M = 3.89$, $SD = 0.871$) and easy to integrate into existing lesson plans ($M = 3.80$, $SD = 0.965$). While participants generally perceived AI as not requiring extensive effort, the item "Using AI in science lessons does not require a lot of effort" had a slightly lower mean score ($M = 3.52$, $SD = 1.094$), indicating some reservations about the effort involved. Similarly, teachers believed that they could use AI tools without needing extensive technical knowledge ($M = 3.69$, $SD = 1.141$). The one-sample t-test results showed that all mean scores were significantly higher than the midpoint, as indicated by statistically significant t-values ($p < .001$), suggesting that teachers' perceptions of AI's ease of use are meaningfully positive.

Table 4: Descriptive Statistics on Attitude Toward AI (AT)

	Mean	SD	T-value	P-value
I am optimistic about using AI in science lessons.	3.92	.890	37.055	< .001
Using AI in science teaching is enjoyable.	4.10	.949	35.657	< .001
I believe that AI is a valuable tool for improving science education.	3.85	1.009	32.103	< .001
I am enthusiastic about incorporating AI into my teaching practice.	3.83	1.014	31.835	< .001
I have a favourable opinion on AI applications in science classrooms.	3.83	1.000	32.287	< .001
Using AI enhances my motivation to teach science effectively.	3.86	1.046	31.087	< .001

Source: Author (2026)

The findings show that teachers generally have a positive attitude toward the use of AI in science teaching. All six items recorded mean scores above the neutral midpoint of 3.0, indicating favourable perceptions. Specifically, teachers reported positive feelings about using AI in their science lessons ($M = 3.92$, $SD = 0.890$) and found it enjoyable to use in teaching ($M = 4.10$, $SD = 0.949$). They also believed that AI was a valuable tool for improving science education ($M = 3.85$, $SD = 1.009$) and expressed enthusiasm for incorporating AI into their teaching practice ($M = 3.83$, $SD = 1.014$). Additionally, participants held favourable views of AI applications in science classrooms ($M = 3.83$, $SD = 1.000$) and agreed that AI enhanced their motivation to teach science effectively ($M = 3.86$, $SD = 1.046$). One-sample t-test results indicated that all mean scores were significantly higher than the midpoint, with all p-values $< .001$, suggesting that positive attitudes toward AI were statistically significant rather than due to chance.

Table 5: Descriptive Statistics on Perceived Behavioural Control (PBC)

	Mean	SD	T-value	P-value
I am confident in my ability to use AI in science teaching.	3.83	1.000	32.287	< .001
I can overcome the technical challenges of using AI in lessons.	3.82	1.019	31.576	< .001
I have sufficient skills to use AI tools effectively in teaching.	3.76	1.021	31.044	< .001
I can integrate AI into my lesson plan.	3.79	.99	31.963	< .001
I can use AI independently without external assistance.	3.73	1.207	26.064	< .001
I can manage classroom activities effectively using AI.	3.70	1.074	29.054	< .001

Source: Author (2026)

The results for perceived behavioural control indicate that teachers generally feel capable of using AI in science teaching. All items scored above the neutral midpoint of 3.0, suggesting that participants perceived they had sufficient control and competence to integrate AI tools into their teaching. Teachers reported confidence in their ability to use AI in science teaching ($M = 3.83$, $SD = 1.000$) and believed they could overcome technical challenges when using AI in lessons ($M = 3.82$, $SD = 1.019$). Participants also indicated that they possessed sufficient skills to use AI tools effectively ($M = 3.76$, $SD = 1.021$) and felt capable of integrating AI into lesson plans ($M = 3.79$, $SD = 0.999$). However, slightly lower mean scores were observed for items related to independent use and classroom management while using AI, with participants indicating moderate confidence in using AI without external assistance ($M = 3.73$, $SD = 1.207$) and managing classroom activities effectively while using AI ($M = 3.70$, $SD = 1.074$). One-sample t-test results showed that all mean scores were significantly higher than the midpoint ($p < .001$), indicating that teachers' perceived behavioural control regarding AI integration was significantly positive.

Table 6: Descriptive Statistics on Behavioural Intention (BI)

	Mean	SD	T-value	P-value
I intend to use AI in science lessons soon.	3.80	1.023	31.326	< .001
I plan to integrate AI tools into my teaching practice regularly.	3.92	1.052	31.357	< .001
I will use AI to enhance student learning.	3.94	1.040	31.939	< .001
I am determined to adopt AI in my teaching practice.	3.72	1.209	25.909	< .001
I intend to explore new AI applications for teaching science.	3.77	1.085	29.316	< .001
I encourage other teachers to use AI in science education.	3.66	1.09	28.187	< .001

Source: Author (2026)

The findings indicate that teachers generally intend to use AI in their science teaching soon. All six items recorded mean scores above the neutral midpoint of 3.0, suggesting positive behavioural intentions toward AI adoption. Participants expressed intentions to regularly integrate AI tools into their teaching practice ($M = 3.92$, $SD = 1.052$) and try to use AI to enhance student learning ($M = 3.94$, $SD = 1.040$). They also indicated that they intend to use AI in their science lessons in the near future ($M = 3.80$, $SD = 1.023$) and plan to explore new AI applications for teaching science ($M = 3.77$, $SD = 1.085$). However, slightly lower mean scores were observed for the determination to adopt AI ($M = 3.72$, $SD = 1.209$) and the intention to encourage other teachers to use AI ($M = 3.66$, $SD = 1.095$), indicating some hesitation in fully advocating AI use among peers. One-sample t-test results revealed that all mean scores were significantly higher than the midpoint ($p < .001$), indicating that teachers' behavioural intentions to adopt AI are statistically significant and not due to chance.

Table 7: Descriptive Statistics on Actual Use (AU)

	Mean	SD	T-value	P-value
I used AI tools in my science lessons during the past semester.	3.70	1.188	26.274	< .001
I used AI applications to prepare teaching materials for science lessons.	3.75	.952	33.161	< .001
I have applied AI-based activities in classroom teaching.	3.63	1.137	26.932	< .001

I used AI tools to assess and monitor student learning outcomes.	3.52	1.119	26.504	< .001
I frequently use AI to enhance my students' engagement and understanding.	3.48	1.286	22.799	< .001

Source: Author (2026)

The results on the actual use of AI in science teaching indicate that teachers have been moderately engaged with AI tools in their instructional practice. All items had scores above the midpoint of 3.0, indicating that participants have used AI to some extent. Teachers reported using AI tools in their science lessons during the past semester ($M = 3.70$, $SD = 1.188$) and utilising AI applications to prepare teaching materials ($M = 3.75$, $SD = 0.952$). They also indicated that they have applied AI-based activities during classroom teaching ($M = 3.63$, $SD = 1.137$) and used AI tools to assess or monitor student learning outcomes ($M = 3.52$, $SD = 1.119$). However, the lowest mean score was observed for the frequency of AI use to enhance student engagement and understanding ($M = 3.48$, $SD = 1.286$), suggesting that consistent or frequent use of AI in classroom instruction may still be limited. One-sample t-test results showed that all mean scores were significantly higher than the midpoint ($p < .001$), indicating that teachers' actual use of AI is significantly positive, although not very high.

4.2. Discussion

This study explored how Ghanaian pre-service science teachers perceive and utilise artificial intelligence (AI) tools in their teaching and lesson preparation, focusing on key constructs such as performance expectancy, effort expectancy, attitude, perceived behavioural control, behavioural intention, and actual use. The findings reveal a generally positive orientation toward AI, with widespread adoption primarily for research, content explanation, and lesson planning rather than assessment or student engagement. These findings align with previous Ghanaian studies indicating that teachers are generally receptive to innovative educational technologies when they perceive clear instructional benefits and opportunities to enhance teaching effectiveness (Asoma et al., 2024; Yeboah et al., 2025). Teachers held favourable views of AI's potential to enhance teaching quality, instructional efficiency, and student learning outcomes, and demonstrated moderate confidence in their ability to use AI effectively. These findings align with the Unified Theory of Acceptance and Use of Technology (UTAUT), which posits that performance expectancy and effort expectancy are among the strongest predictors of technology adoption intentions (Venkatesh et al., 2003). The positive perceptions observed in this study suggest that pre-service science teachers recognise the pedagogical value of AI and believe that it can improve teaching effectiveness. Similar findings have been reported among Ghanaian teachers, where positive perceptions of technology initiatives significantly influenced classroom technology integration (Asoma et al., 2024). The findings also align with evidence from broader African contexts, where educators increasingly view AI as a tool to improve instructional delivery, learner engagement, and access to educational resources (Adigun et al., 2025; Akanzire et al., 2025). However, despite strong behavioural intentions to use AI, actual classroom integration remained moderate, suggesting that AI adoption among Ghanaian pre-service science teachers is still in its developmental stage.

The interpretation of these results suggests that while pre-service teachers recognise the instrumental value of AI as a cognitive and preparatory aid, its transformative pedagogical potential has yet to be fully realised. The predominance of ChatGPT and other generative AI tools for research and content development aligns with an instrumentalist view of technology, in which AI serves to augment teachers' efficiency rather than fundamentally reshape instructional methods. This pattern is consistent with the Unified Theory of Acceptance and Use of Technology (UTAUT), in which performance expectancy strongly predicts adoption intentions (Venkatesh et al., 2003).

The positive effort expectancy scores indicate that respondents generally perceived AI tools as accessible, understandable, and relatively easy to integrate into science teaching. Within the UTAUT framework, effort expectancy influences behavioural intention by reducing the cognitive burden associated with learning and applying new technologies (Venkatesh et al., 2003). The findings suggest that the increasing availability of user-friendly AI applications, such as ChatGPT, has lowered the barriers to technology adoption among pre-service teachers. Similar observations have been reported among educators in Ghana and other developing countries, where ease of use significantly enhances willingness to adopt emerging technologies (Agyei & Voogt, 2010; Naatu

et al., 2024). Nevertheless, the slightly lower ratings regarding independent use and technical competence indicate that some participants may still require additional training and institutional support to maximise the pedagogical benefits of AI. This finding is consistent with evidence suggesting that personal dispositions and affective competencies influence educational decision-making and engagement with innovative practices among pre-service teachers (Mintah et al., 2020; Budu et al., 2022). The favourable attitudes toward AI and the positive levels of perceived behavioural control observed among participants provide further support for the Theory of Planned Behaviour (TPB) (Ajzen, 1991). According to the TPB, behavioural intention is shaped by individuals' attitudes toward a behaviour and their perceived capacity to perform it successfully. The positive attitudes recorded in this study suggest that pre-service science teachers perceive AI as a valuable instructional resource that can enhance science teaching and learning. Likewise, their confidence in overcoming technical challenges and integrating AI into lesson planning reflects relatively high perceived behavioural control. These findings imply that participants possess both the motivation and self-belief necessary for AI adoption. Similar patterns have been reported among Ghanaian pre-service teachers, where confidence, emotional competence, and personal dispositions significantly influenced engagement with innovative educational practices (Mintah et al., 2020; Budu et al., 2022). Consequently, the findings reinforce TPB assumptions that favourable attitudes and strong perceptions of control are important precursors to behavioural intentions to adopt technology. However, the intention-behavior gap evident in actual use highlights systemic and infrastructural constraints (Kim, 2025). These may include limited access to reliable digital infrastructure, insufficient professional development, and unclear institutional policies, which collectively impede the translation of positive dispositions into sustained classroom practice (Hakimi & Shahidzay, 2024; Kim & Wargo, 2025).

These findings resonate with the broader international literature documenting early-stage AI adoption in education as predominantly preparatory and supportive rather than transformative. Similar studies have identified performance expectancy and ease of use as pivotal predictors of AI acceptance among educators, reinforcing the external validity of these models within the Ghanaian context (Geddam et al., 2024).

The relatively lower use of AI for assessment and student engagement highlights an important contextual dimension of technology adoption in developing countries. While educators recognised the benefits of AI for lesson preparation, research, and content explanation, they appeared less confident in employing AI for more sophisticated pedagogical functions. This finding reflects broader patterns reported across Sub-Saharan Africa, where limitations in digital infrastructure, institutional support, and professional development often constrain the depth of technology integration (Hakimi & Shahidzay, 2024; Naatu et al., 2024). In Ghana specifically, previous studies have shown that positive perceptions alone do not guarantee effective classroom implementation unless accompanied by adequate technological resources, training opportunities, and supportive educational policies (Abedi et al., 2023; Asoma et al., 2024). Consequently, the gap between intention and actual use observed in this study may be explained not only by individual factors but also by systemic and contextual barriers that shape technology adoption behaviour.

The implications of these findings are significant for Ghana's science education system, which appears to be at a critical juncture. The widespread positive perceptions and intentions indicate that pre-service teachers are ready to embrace AI, suggesting fertile ground for targeted interventions to deepen AI integration (Abedi et al., 2023; Agyei & Voogt, 2010; Blonder et al., 2024). Addressing systemic barriers could facilitate the transition from instrumental use toward more transformative pedagogical practices. Similar transformations have been observed in science classrooms where collaborative and learner-centred instructional strategies promoted deeper conceptual understanding and more meaningful student engagement (Amponsah & Ochonogor, 2018). These unlock AI's full potential to enhance inquiry-based learning and student engagement in science education.

Theoretically, this study reinforces integrative adoption frameworks that combine expectancy-based constructs with behavioural and contextual factors, confirming that perceptual drivers are necessary but not sufficient for educational innovation. Practically, it signals to policymakers and educational leaders that investments in digital infrastructure, curriculum alignment, and sustained teacher training are essential to capitalise on the current momentum and ensure equitable and effective AI adoption (Mustofa et al., 2025). In summary, this research contributes to the growing body of knowledge on AI in education by providing empirical evidence from a Sub-

Saharan African context and emphasising the interplay between individual perceptions and systemic enablers in shaping AI integration trajectories.

The findings suggest that an interplay of expectancy beliefs, behavioural dispositions, and contextual conditions influences AI adoption among Ghanaian pre-service science teachers. Consistent with UTAUT, perceptions of usefulness and ease of use contributed to positive intentions toward AI adoption. At the same time, TPB constructs demonstrated the importance of attitudes and perceived behavioural control in shaping behavioural intentions. However, the persistence of an intention–practice gap indicates that psychological readiness alone is insufficient to ensure sustained AI integration. The findings therefore support contemporary African scholarship that advocates a holistic approach to educational technology adoption, combining individual capacity building with institutional investment, policy support, and improved digital infrastructure to facilitate meaningful and sustainable AI integration in teaching and learning.

5. Conclusion

This study demonstrates that Ghanaian pre-service science teachers generally hold positive perceptions of artificial intelligence, recognise its usefulness, and are motivated to integrate it into teaching, particularly for research, lesson planning, and content explanation; yet, actual classroom use remains moderate. These findings indicate a transitional phase in science education, in which cognitive and attitudinal readiness is high but systemic and contextual factors, such as technological infrastructure, professional development, and curriculum alignment, moderate full integration. The study suggests that policies and interventions should prioritise capacity building, practical training, and supportive structures to bridge the gap between intention and practice, thereby enabling AI to move from a supportive tool to a transformative educational resource. Future research should investigate longitudinal adoption, pedagogical impacts on student outcomes, and regional variations, highlighting that realising AI's potential in Ghanaian science classrooms depends on both teacher readiness and systemic facilitation.

5.1 Limitation of the Study

A key limitation of this study is its reliance on self-reported data from pre-service science teachers, which may be influenced by social desirability bias or overestimation of AI use and competence. Additionally, the cross-sectional design captures perceptions and behaviours at a single point in time, limiting the ability to infer causal relationships or track changes in AI adoption over time. The study also focused primarily on teachers' perspectives without incorporating insights from students or school administrators, which may influence AI integration. Finally, the findings are specific to the Ghanaian context. While they provide valuable insights for similar educational settings, generalisation to other countries with different infrastructural and policy conditions should be approached with caution.

Author Contributions: All authors contributed to this research.

Conflict of Interest declaration: The authors declare that they have no competing interests or conflicts of interest related to this study.

Data Access Statement: Research data supporting this publication are available upon request.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Abedi, E. A., Prestridge, S., & Hodge, S. (2023). Teachers' beliefs about technology integration in Ghana: A qualitative study of teachers', headteachers', and education officials' perceptions. *Education and Information Technologies*, 29(5), 5857–5877. <https://doi.org/10.1007/s10639-023-12049-0>
- Adigun, O. T., Tijani, F. A., Haihambo, C. K., & Enock, S. L. (2025). Understanding pre-service teachers' intention to adopt and use artificial intelligence in inclusive classrooms in Nigeria. *Frontiers in Education*, 10. <https://doi.org/10.3389/educ.2025.1519472>
- Agormedah, E. K., Ankomah, F., Frimpong, J. B., Quansah, F., Srem-Sai, M., Hagan, J. E., & Schack, T. (2022). Investigating teachers' experience and self-efficacy beliefs across gender in implementing the new standards-based curriculum in Ghana. *Frontiers in Education*, 7. <https://doi.org/10.3389/educ.2022.932447>
- Agyei, D. D., & Voogt, J. (2010). ICT use in the teaching of mathematics: Implications for professional development of pre-service teachers in Ghana. *Education and Information Technologies*, 16(4), 423–439. <https://doi.org/10.1007/s10639-010-9141-9>
- Ajzen, I. (1991). The theory of planned behaviour. *Organisational Behaviour and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I., & Schmidt, P. (2020). Changing behaviour using the theory of planned behaviour. In M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen, & T. Lintunen (Eds.), *The handbook of behaviour change* (pp. 17–31). Cambridge University Press.
- Akanzire, B. N., Nyaaba, M., & Nabang, M. (2025). Generative AI in teacher education: Teacher educators' perception and preparedness. *Journal of Digital Educational Technology*, 5(1), ep2508. <https://doi.org/10.30935/jdet/15887>
- Akyeampong, K. (2017). Teacher educators' practices and visions of good teaching in the context of teacher education reform in Ghana. *Educational Researcher*, 46(4), 194–203. <https://doi.org/10.3102/0013189X17711907>
- Alghamdy, R. Z. (2023). Pedagogical and ethical implications of artificial intelligence in the EFL context: A review study. *English Language Teaching*, 16(10), 87–97. <https://doi.org/10.5539/elt.v16n10p87>
- Amponsah, K. D., & Ochonogor, C. E. (2018). Facilitating conceptual change in students' understanding of electrochemistry through a collaborative teaching strategy. *American Journal of Educational Research*, 6(6), 596–601. <https://doi.org/10.12691/education-6-6-2>
- Asoma, R. K., Agyei, D. D., Assamah, G., & Amponsah, K. D. (2024). Examining teachers' perceptions of the impact of the Government of Ghana's Wi-Fi technology program on teaching practices: An empirical study from senior high schools in the Greater Accra Region. *Cogent Education*, 11(1), Article 2391116. <https://doi.org/10.1080/2331186X.2024.2391116>
- Bayaga, A., & du Plessis, A. (2024). Ramifications of the Unified Theory of Acceptance and Use of Technology (UTAUT) among developing countries' higher education staff. *Education and Information Technologies*, 29(8), 9689–9714.
- Blonder, R., Feldman-Maggor, Y., & Rap, S. (2024). Are they ready to teach? Generative AI as a means to uncover pre-service science teachers' PCK and enhance their preparation program. *Journal of Science Education and Technology*, 34(6), 1301–1310. <https://doi.org/10.1007/s10956-024-10180-2>
- Bosnjak, M., Ajzen, I., & Schmidt, P. (2020). The theory of planned behaviour: Selected recent advances and applications. *Europe's Journal of Psychology*, 16(3), 352–356.
- Budu, G. B., Kemetse, J. K., & Amponsah, K. D. (2022). A Ghanaian study on multiple intelligences of pre-service science teachers in selected colleges of education. *Education Quarterly Reviews*, 5(1), 15–27. <https://doi.org/10.31014/aior.1993.05.01.418>
- Budu, K. W. A., Yinping, M., & Mireku, K. K. (2018). Investigating the effect of behavioural intention on e-learning system usage: An empirical study of tertiary education institutions in Ghana. *Mediterranean Journal of Social Sciences*, 9(3), 201–216.
- Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M. I., & Andrade-Abarca, P. S. (2024). The impact of pedagogical beliefs on the adoption of generative AI in higher education: Predictive model from UTAUT2. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/fraci.2024.1497705>
- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Dignum, V. (2023). Responsible artificial intelligence: Recommendations and lessons learned. In *Responsible Artificial Intelligence* (pp. 195–214). Springer. https://doi.org/10.1007/978-3-031-08215-3_9
- Eusebio, E. J. G., Baldera, P., Patiam, A. M., Villanueva, E. R., Gaa, N. A., Solis, A. M. F., Soriano, M. L. C., & Ribon, A. L. (2025). AI in the classroom: A systematic review of barriers to educator acceptance.

- International Journal of Learning, Teaching and Educational Research*, 24(9), 126–147. <https://doi.org/10.26803/ijlter.24.9.7>
- Fang, W., Na, M., & Alam, S. S. (2025). Usage intention of AI among academic librarians in China: Extension of the UTAUT model. *Sustainability*, 17(7), 2833. <https://doi.org/10.3390/su17072833>
- Geddam, S. M., Nethravathi, N., & Hussian, A. A. (2024). Understanding AI adoption: The mediating role of attitude in user acceptance. *Journal of Informatics Education and Research*, 4(2), 1664–1672. <https://doi.org/10.52783/jier.v4i2.975>
- Hakimi, M., & Shahidzay, A. K. (2024). Transforming education with artificial intelligence: Potential and obstacles in developing countries. *Preprints*. <https://doi.org/10.20944/preprints202407.2542.v1>
- Info, A., Agyemang, M., Amponsah, S., Van Wyk, M. M., Huang, R., Tlili, A., Hosny, A., Metwally, S., Huanhuan, G. G., Adarkwah, A., Shehata, B., & Wang, H. (2023). Awareness and acceptance of ChatGPT as a generative conversational AI for transforming education by Ghanaian academics: A two-phase study. *Journal of Applied Learning & Teaching*, 6(2). <https://doi.org/10.37074/jalt.2023.6.2.26>
- Iqbal, S., & Ahmed Bhatti, Z. (2015). An investigation of university student readiness towards m-learning using the technology acceptance model. *International Review of Research in Open and Distributed Learning*, 16(4). <https://doi.org/10.19173/irrodl.v16i4.2351>
- Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A. A., Hattab, M. K., & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, 14(10), 1090. <https://doi.org/10.3390/educsci14101090>
- Kim, J. (2025). Perceptions and preparedness of K–12 educators in adopting generative AI. *Research in Learning Technology*, 33. <https://doi.org/10.25304/rlt.v33.3448>
- Kim, J., & Wargo, E. (2025). Empowering educational leaders for AI integration in rural STEM education: Challenges and strategies. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1567698>
- Kim, Y., Blazquez, V., & Oh, T. (2024). Determinants of generative AI system adoption and usage behaviour in Korean companies: Applying the UTAUT model. *Behavioral Sciences*, 14(11), 1035. <https://doi.org/10.3390/bs14111035>
- Mustafa, A. N. (2024). The future of mathematics education: Adaptive learning technologies and artificial intelligence. *International Journal of Science and Research Archive*, 12(1), 2594–2599. <https://doi.org/10.30574/ijrsra.2024.12.1.1134>
- Mustofa, H. A., Kola, A. J., & Owusu-Darko, I. (2025). Integration of artificial intelligence (ChatGPT) into science teaching and learning. *International Journal of Ethnoscience and Technology in Education*, 2(1), 108. <https://doi.org/10.33394/ijete.v2i1.14195>
- Naatu, F., Selormey, F. S., & Naatu, S. (2024). Determinants of digital technology adoption in sub-Saharan Africa: Ghana. *International Journal of Emerging Markets*, 20(10), 4111–4133. <https://doi.org/10.1108/IJOEM-09-2023-1503>
- Naskar, S. T., & Lindahl, J. M. M. (2025). Forty years of the theory of planned behaviour: A bibliometric analysis (1985–2024). *Management Review Quarterly*, 1–60.
- Ofosu-Ampong, K. (2023). Gender differences in perception of artificial intelligence-based tools. *Journal of Digital Art & Humanities*, 4(2), 52–56. https://doi.org/10.33847/2712-8149.4.2_6
- Ofosu-Ampong, K., Acheampong, B., Kevor, M. O., & Amankwah-Sarfo, F. (2023). Acceptance of artificial intelligence (ChatGPT) in education: Trust, innovativeness, and students' psychological needs. *Information and Knowledge Management*, 13(4). <https://doi.org/10.7176/IKM/13-4-03>
- Oyewole, O. (2018). Performance expectancy, effort expectancy, and facilitating conditions as factors influencing smartphone use for mobile learning by postgraduate students of the University of Ibadan, Nigeria. *Interdisciplinary Journal of e-Skills and Lifelong Learning*, 14, 95–115.
- Roshan, S., Zaffar Iqbal, S., & Qing, Z. (2024). Teacher training and professional development for implementing AI-based educational tools. *Journal of Asian Development Studies*, 13(2), 1972–1987. <https://doi.org/10.62345/jads.2024.13.2.154>
- Sova, R., Tudor, C., Tartavulea, C. V., & Diaconescu, R. I. (2024). Artificial intelligence tool adoption in higher education: A structural equation modelling approach to understanding impact factors among economics students. *Electronics*, 13(18), 3632. <https://doi.org/10.3390/electronics13183632>
- Teo, T., & Noyes, J. (2012). Explaining the intention to use technology among pre-service teachers: A multi-group analysis of the Unified Theory of Acceptance and Use of Technology. *Interactive Learning Environments*, 22(1), 51–66. <https://doi.org/10.1080/10494820.2011.641674>
- Tram, N. H. M. (2024). Unveiling the drivers of AI integration among language teachers: Integrating UTAUT and AI-TPACK. *Computers in the Schools*, 42(2), 100–120. <https://doi.org/10.1080/07380569.2024.2441155>
- Venkatesh, V., & Thong, J. Y. L. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.

Yeboah, R., Amponsah, K. D., Commey-Mintah, P., Sedofia, J., & Donkor (Opare), P. B. K. (2025). Game-based learning in Ghanaian primary schools: Listening to teachers' views. *Education 3-13*, 53(1), 99–113. <https://doi.org/10.1080/03004279.2023.2290903>

Which Decisions Remain the Educator's Own: The Hybrid Decision-Making Model for AI-Integrated Education

Pauline P. L. Chin¹

¹ Meragang Sixth Form College
ORCID ID: <https://orcid.org/0000-0001-5635-9141>

Abstract

When educators working with artificial intelligence (AI) are asked which decisions remain their own to make, most know the boundary exists. Few have a structure for it. This paper addresses that absence directly by proposing the Hybrid Decision-Making Model, a three-tier framework that assigns decision authority across AI-integrated higher education practice. Tier 1 covers routine, reversible decisions where AI leads and the human retains oversight. Tier 2 covers decisions requiring both AI-provided data and human contextual judgement, where neither contribution alone produces a valid outcome. Tier 3 covers decisions whose irreversible consequences, ethical weight, and relational specificity make human authority a structural requirement rather than a preference. The HDM is grounded in three theoretical foundations: Simon's (1947, 1990) account of bounded rationality, which explains why the absence of decision authority produces predictable drift towards AI-generated outputs; Engelbart's (1962) and Baker's (2016) principle of augmentation, which establishes that AI should extend human capability rather than replace it; and the hybrid intelligence and governance frameworks of Holstein et al. (2019, 2020), Akata et al. (2020), and IMDA and PDPC (2020), which provide the dimensions and human involvement taxonomy from which the three tiers are derived. Empirical support is drawn from a qualitative study of seven higher education educators who consistently drew the same structural distinctions the HDM formalises, in the absence of a formally named decision authority structure for AI-integrated practice. The model does not prescribe institutional behaviour. It names the conditions under which decision authority is maintained. The structure was already there. This paper simply gives it a name.

Keywords: Artificial Intelligence, Higher Education, Decision Authority, Human-AI Collaboration, Hybrid Decision-Making, AI Governance

1. Introduction

A science educator, when asked about artificial intelligence in the classroom, did not question whether the technology belonged there. The question that arose was more precise: which decisions remained the educator's own to make. This finding, drawn from a qualitative study of seven higher education educators (Chin, 2025), points to a structural problem that enthusiasm for AI integration has so far outpaced. A review of 24 national AI policy strategies found that the use of AI in education is largely absent from policy conversations, whilst the instrumental value of education in producing an AI-ready workforce is overwhelmingly prioritised (Schiff, 2022). Most educators working with AI today would recognise this clearly: they know the boundary between human and machine decision-making exists, they feel it every time a choice must be made, but they do not yet have a map for it. Higher education institutions are therefore adopting AI tools without an agreed structure for determining which

decisions AI should lead, which educators should lead, and which require the contribution of both. Even national governance frameworks that have begun to address AI in organisational settings have not resolved this question at the level of classroom practice (AITI, 2025; IMDA and PDPC, 2020). The problem is not one of willingness. It is one of design.

The consequences of this absence are already visible, though rarely named. When institutional guidance on artificial intelligence is unclear, educators do not stop; they adapt, drawing their own boundaries as they go. This produces inconsistency across classrooms (Chan, 2023). Without a shared structure, decisions made under time pressure tend naturally towards convenience. Pedagogical judgement shifts gradually as a result: the question moves from what does this student need to what does the AI suggest, and the two are not always the same. Students, meanwhile, receive different messages about acceptable use depending on which classroom they are in, which produces uneven habits and an uneven relationship with their own thinking. Uncertainty about academic integrity follows, not from dishonesty on anyone's part, but from the absence of clear and shared expectations. Moments of genuine good practice occur across institutions, but without a common structure they remain individual rather than collective achievements. The difficulty, therefore, is not that educators or institutions have failed. It is that the situation has moved faster than the frameworks available to manage it.

The research field has made substantial progress in identifying the tensions that AI integration introduces into higher education. Studies have documented the risk that over-reliance on AI-generated material weakens students' capacity for independent thinking (Chin, 2025; Popenici and Kerr, 2017), that academic integrity becomes uncertain when the authorship of student work is unclear (Chin, 2025; Chan, 2023), and that data privacy and ethical responsibility remain inadequately resolved in most institutional contexts (Chan, 2023; Floridi and Cows, 2019; Miao and Holmes, 2021). The potential for AI systems to produce biased or inaccurate outputs has been noted as a concern that existing safeguards have not yet addressed sufficiently (Chan, 2023), whilst unequal access to AI tools risks widening educational disparities that institutions are already working to close (Schiff, 2022; Miao and Holmes, 2021). The governance challenges that arise when AI is integrated into institutional decision-making processes have been documented in regional policy contexts, including within the Southeast Asian setting in which this paper is situated (Yar et al., 2024). The field has also examined educator attitudes, student behaviour, and policy developments in AI adoption with increasing depth and careful development (Chin, 2025; Chan, 2023; Schiff, 2022). A scoping review of 32 empirical studies confirmed that balancing AI-assisted and human-centred assessment remains an unresolved challenge, with educators calling for clearer boundaries around what AI can and cannot determine in student work (Xia et al., 2024). In this sense, the problem is well understood. What the literature has not yet produced is a practical structure for decision-making: a clear account of which decisions artificial intelligence may lead, which must remain under human judgement, and which should involve both.

That absence has a common source. Each risk identified in the literature points to the same underlying issue: the absence of clear control over decision-making in AI-integrated practice. The risks include weakened independent thinking (Chin, 2025; Popenici & Kerr, 2017), uncertain academic integrity (Chin, 2025; Chan, 2023), unresolved ethical responsibility (Chan, 2023; Floridi & Cows, 2019), unchecked AI outputs (Chan, 2023), and unequal access (Schiff, 2022; Miao & Holmes, 2021). In each case, the difficulty arises not from the presence of artificial intelligence itself, but from the absence of guidance on who should decide, when that decision should be made, and on what basis.

It is the quiet centre of all the literature has been circling. The field has mapped where control is missing, even when it has not named the mapping as such. This is not simply a practical difficulty. Simon (1947, 1990) demonstrated that human decision-making operates under real cognitive constraints: individuals do not optimise across all available choices but settle for what is sufficient given the time, information, and capacity available. Without a deliberate structure to direct judgement, decisions made under pressure default to convenience rather than principle. The absence of control is not merely inconvenient. Simon (1947, 1990) shows it is structurally dangerous.

Three bodies of work established the conditions that made a response to this danger both necessary and possible. Baker (2016) rejected full automation, demonstrating that AI should extend human capability rather than replace

it and that the human must remain present in decision-making. Holstein et al. (2019, 2020) established that human-AI collaboration is not natural or automatic. It must be deliberately designed, and roles cannot be left vague. Akata et al. (2020) identified the distribution of decision-making between human and machine as the central unresolved problem in hybrid intelligence research. Each contribution advanced the field. None produced the structure itself. Each described how work should be distributed between human and machine. None specified who holds authority when a judgement must be made.

This study addresses that directly. Drawing on semi-structured interviews with seven experienced higher education educators, it examines how these educators drew decision-making boundaries in their own teaching, in the absence of formal institutional guidance (Chin, 2025). The study is small and qualitative by design. Its value lies not in breadth but in the depth of seven educators' lived experiences, which reveal what large-scale surveys cannot: the precise distinctions practitioners make when no framework exists to guide them. The central question the study asks is this: how should decision-making authority be distributed between artificial intelligence and human educators in higher education practice? What the study reveals is that educators were already making these distinctions in practice, intuitively and consistently, long before a formal structure existed to support them. The Hybrid Decision-Making Model does not invent a new structure. It names and formalises one that practitioners have already been living.

The paper proceeds as follows. Section 2 reviews the three bodies of literature from which the HDM is constructed: hybrid intelligence frameworks, teacher-AI collaboration research, and the division of labour between human and machine. Section 3 establishes the theoretical foundations of the model. Section 4 presents the Hybrid Decision-Making Model in full. Section 5 draws on the empirical study to validate the framework across its three tiers. Section 6 offers reflections on how the model might inform institutional practice. Section 7 discusses the paper's contributions in relation to the field. Section 8 identifies directions for future research. The science educator whose question opened this paper asked which decisions remained their own to make. This paper answers that question. Not with a reassurance, but with a structure.

2. Literature Review

Three bodies of scholarship inform the theoretical foundations of the Hybrid Decision-Making Model. The first examines hybrid intelligence in education. The second examines teacher-AI collaboration. The third examines the division of labour between human and machine. Each advances the field substantially. Each stops at the same point: no study identified in this review specifies how decision-making authority should be distributed in practice.

2.1. Hybrid Intelligence in Education

Akata et al. (2020) established that combining human and artificial capability effectively requires systems that are collaborative, adaptive, responsible, and explainable, and identified the distribution of decision-making between human and machine as a problem the field had not yet resolved. Bredeweg and Kragten (2022) demonstrated through a case study in secondary education that intelligent tutoring systems work more effectively when teachers remain involved, not by doing what the system does alongside it, but by focusing on what it cannot do: interpreting data about student progress and making pedagogical judgements that require human experience.

Holstein et al. (2020) identified four dimensions structuring human and AI instruction: instructional goals, relevant information, instructional actions, and decision-making processes. Their contribution is showing that these dimensions exist and can be mapped. What the framework does not address is who holds authority at each dimension. It describes how human and AI instruction can be combined but does not specify which actor makes the final decision at each point, or on what basis.

2.2. Teacher-AI Collaboration

Baker (2016) argued that rather than building AI systems capable of replacing teacher decision-making, the field should focus on building systems that provide teachers with better data and analysis, leaving the decisions

themselves with the educator. Baker named this intelligence amplification. The human decides. The AI provides the information on which that decision is based.

Holstein et al. (2019) investigated what intelligence amplification looks like in practice through a study of teacher and student needs in AI-enhanced classrooms. They found that teachers wanted real-time data from AI systems to support their decisions, not automated decisions made on their behalf. The design of collaboration must be deliberate. What Holstein et al. (2019) do not specify is the decision authority structure that design should produce.

Paiva and Bittencourt (2020) confirmed that this collaborative model is achievable in real settings. Their T-Partner authoring tool processes educational data to support instructors in making pedagogical decisions without making those decisions for them. This confirms that the model Baker proposed holds under real working conditions.

2.3 The Division of Labour Between Human and Machine

Engelbart (1962) argued that computing technology exists to augment human intellect rather than automate it. That principle sits at the foundation of the HDM.

Chou, Huang and Lin (2011) applied this principle in a higher education tutoring context, dividing tutoring responsibilities between software and teacher. The software evaluated answers and generated hints. The teacher interpreted student thinking and made relational decisions. Their findings confirmed that this division reduced teacher workload without reducing student support quality. The main condition for success was that each party's responsibilities were defined clearly from the outset.

Holstein and Alevan (2022) confirmed the same pattern in K-12 classrooms through a study involving Lumilo, a tool displaying real-time student learning data on smart glasses worn by the teacher. Students learned more when the teacher used Lumilo alongside the AI tutoring system because the teacher remained the decision-maker: Lumilo provided the information and the teacher decided how to respond. None of these studies, however, specifies who holds decision authority when a judgement call is required. Allocating tasks is not the same as assigning decision authority. The HDM addresses that distinction.

2.4 The Common Gap: Distributing Work Without Distributing Decisions

There is a precise distinction between distributing work and distributing decisions. Distributing work assigns tasks. Distributing decisions assigns authority: when a judgement is required, whose call is it? A teacher can be assigned the task of reviewing AI-generated feedback without it being established whether that teacher has the authority to change it, question its basis, or set it aside. The task has been allocated. The authority has not.

The consequences of leaving decision authority unspecified are not theoretical. Simon (1947, 1990) demonstrated that without a structure that clearly assigns decision authority, an educator working through a full teaching day will not pause to consider whether a particular decision is theirs to make. The default will be convenience. Over time, convenience quietly replaces principle. Decisions drift towards the AI. Neither drift is dramatic. Neither is immediately visible. Both are consequential.

Akata et al. (2020) identified the distribution of decision-making as the central unresolved problem in hybrid intelligence research. Holstein et al. (2019) established that human-AI collaboration must be deliberately designed but did not specify what the design of decision authority should look like. Baker (2016) argued that the human must remain in the decision loop but did not define which decisions that loop should contain. It is the precise point at which the existing research reaches its limit.

A concurrent contribution moves further into this territory than the frameworks above. Borchers, Viberg, and Kizilcec (2026) introduce the Agency Allocation Framework (AAF), which reframes learner agency as the allocation of decision-making authority across learners, educators, institutions, and AI systems in large-scale, automated learning environments. Their framework treats decision authority, rather than task distribution, as its

central object, and it names four recurring challenges that keep that authority from being studied or designed with precision: conceptual ambiguity, the difficulty of measuring agency at scale, the trade-off between agency and efficiency, and the redistribution of agency through AI mediation. Its diagnosis strengthens the case that a structural resolution is needed. Its locus and its purpose, however, differ from the one this paper proposes. The AAF addresses learner agency in large-scale automated systems, and it stops at naming the challenges that make authority hard to allocate. The HDM addresses educator agency in general AI-integrated higher education practice, and it does not stop at diagnosis. It proposes the tiers themselves, and a test for assigning a decision to one of them. Where the AAF clarifies why decision authority has remained elusive, the HDM specifies where it should sit.

The Hybrid Decision-Making Model begins where that limit ends.

3. Theoretical Foundations

The literature review established three findings that this section addresses directly. First, human-AI collaboration produces better outcomes when deliberately designed (Akata et al., 2020; Holstein et al., 2019). Second, existing frameworks have distributed work with precision but none specifies who holds authority at the moment of judgement (Baker, 2016; Holstein et al., 2020; Paiva and Bittencourt, 2020). Third, in the absence of a clear decision authority structure, cognitive pressure produces consistent and predictable drift (Simon, 1947, 1990). The HDM is the theoretical response to all three findings.

The section proceeds in three stages. Section 3.1 draws on Simon (1947, 1990) to establish why decision authority must be made explicit. Section 3.2 draws on Engelbart (1962) and Baker (2016) to establish augmentation as the governing principle. Section 3.3 derives the HDM's three tiers from Holstein et al. (2019, 2020), Akata et al. (2020), and the human involvement taxonomy of IMDA and PDPC (2020), subsequently adopted by Brunei Darussalam as its national guide on AI governance and ethics (AITI, 2025).

3.1. Bounded Rationality and the Design of Decision Authority

Drift, as used in this paper, refers to a directional and incremental shift away from deliberate human judgement and towards uncritical acceptance of AI-generated outputs. Under the conditions Simon (1947, 1990) identifies, it is predictable.

Simon's account of bounded rationality begins with a structural observation about the limits of human cognition. Decision-making proceeds not by surveying all available options and selecting the best one, but by finding an option that meets a minimum threshold and stopping. Simon called this satisficing: the decision that is good enough, given the information and time available, is the decision that is made. The constraint is cognitive: working memory is finite, attention is depletable, and the number of decisions competing for both across a teaching day is not.

Search is also costly. Reviewing an AI output more carefully, questioning its basis, comparing it against independent professional judgement all require time and resource that are not freely available. Bounded rationality names the limit on decision-making capacity. Search cost names why that limit is not overcome even when the educator is aware of it.

In an AI-integrated classroom without a clear decision authority structure, both mechanisms operate simultaneously. Before a substantive decision can be made, a prior question must first be resolved: whose decision is this? That meta-decision consumes resource before the substantive judgement begins. Under time pressure, it resolves towards the most immediately available option which, in an AI-integrated classroom, is consistently the AI's output.

The result is drift. The AI produces a satisfactory output; the absence of a structural signal that scrutiny is required increases the cost of providing it; the output is accepted; each acceptance recognises the next. The drift is gradual because no single acceptance constitutes a recognisable failure. It is directional because the asymmetry of

cognitive cost consistently favours the AI's output. It is unnoticed because each individual decision is locally defensible. What accumulates is not negligence. It is a structural realignment of where educational judgement is exercised, produced not by intention but by the absence of design.

Awareness alone cannot correct drift. Training addresses what educators know about the problem. It does not change the conditions that produce it. When decision authority is clearly assigned, the educator no longer needs to resolve the prior question of whose decision this is. Cognitive resource that was being spent on the meta-decision becomes available for decisions that genuinely require deliberate human judgement. Distributing work without distributing decisions leaves the meta-decision unresolved at every point of practice. Under the conditions Simon describes, it resolves towards convenience. That is the direction a decision authority structure is designed to interrupt.

3.2. Augmentation as the Governing Principle

If bounded rationality produces drift when decision authority is unclear, the governing principle for distributing that authority must reduce cognitive load without displacing human judgement. Only one arrangement satisfies both conditions: augmentation.

Engelbart (1962) argued that computing technology exists to extend human intellectual capability, not replace it. A technology that replaces human decision-making removes the contribution that gives the decision its validity in complex situations. A technology that extends it preserves that contribution whilst reducing the cognitive burden required to exercise it. Baker (2016) translated this into a concrete decision hierarchy: build systems that provide teachers with better information, leaving the decision with the educator. The human decides. The AI provides the basis on which that decision is made. Augmentation reduces the cognitive cost of deciding without transferring the authority to decide.

Without augmentation, efficiency and scrutiny compete directly, which is the condition Section 3.1 identified as productive of drift. Augmentation dissolves that tension by allowing AI to handle information-processing tasks that consume cognitive resource, whilst preserving the educator's authority over the judgements that require human experience and ethical responsibility. The educator is not burdened by what the AI can do well. The educator is freed for what only the educator can do.

This principle runs through every tier of the HDM. At the task level, augmentation is efficiency with oversight. At the judgement level, it is support without substitution.

3.3. From Hybrid Adaptivity to Decision Authority

The three tiers of the HDM were already latent in the existing literature. This subsection makes them explicit. Holstein et al. (2020) identified four dimensions along which human and AI instruction can be combined: instructional goals, relevant information, instructional actions, and decision-making processes. The framework demonstrates that hybrid adaptivity is achievable. What it does not specify is where authority is held at the moment of judgement.

Akata et al. (2020) established four criteria for productive human-AI systems: collaborative, adaptive, responsible, and explainable. The HDM accepts those four and adds a fifth: clear. A system that does not specify who decides remains structurally incomplete. Without clarity on decision authority, the meta-decision burden Section 3.1 identified is reproduced at every point of practice.

Singapore's Model AI Governance Framework (IMDA and PDPC, 2020), subsequently adopted by Brunei Darussalam (AITI, 2025), distinguished three levels of human involvement in AI-augmented decision-making: the human makes the final decision informed by AI analysis; AI leads whilst the human monitors and retains authority to intervene; and AI operates autonomously with oversight maintained at the system level. These levels were

designed for organisational governance. Translating them into a decision authority structure for higher education classroom practice is what the HDM performs.

That translation requires one additional dimension the governance framework does not supply: decision consequence, referring to the degree to which a decision is reversible, ethically significant, and dependent on contextual and relational knowledge that only human experience provides. When the level of human involvement is mapped against the level of consequence, three distinct and stable configurations emerge: AI leads with human oversight; both contribute and neither decides alone; and human primacy is a structural requirement. Each resolves the meta-decision burden by specifying in advance whose decision this is.

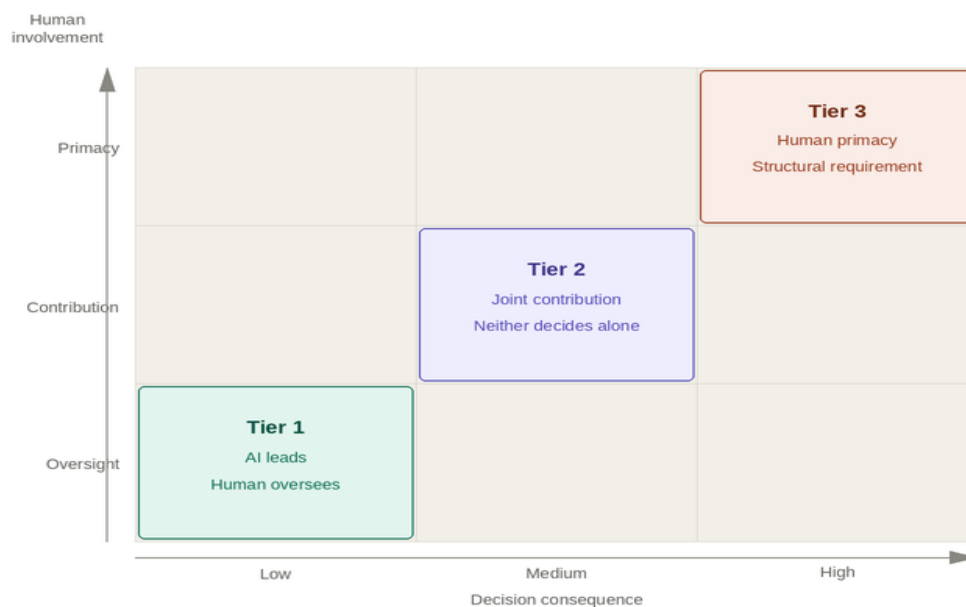


Figure 1: The interaction of human involvement and decision consequence producing three configurations of decision authority in the Hybrid Decision-Making Model.

The distinction between the HDM and its source frameworks is precise. Existing frameworks describe how humans and AI interact. The HDM specifies who holds authority when a judgement must be made. It is a decision authority taxonomy, assigning ownership of judgement across three levels of consequence. Without this resolution, existing frameworks remain descriptive rather than operational. The HDM replaces structural ambiguity with structural clarity. The premises established by Holstein et al. (2020), Akata et al. (2020), and IMDA and PDPC (2020) produce a condition that requires resolution. The HDM is that resolution.

Under conditions of bounded rationality, the absence of decision authority produces predictable drift. Augmentation provides the only governing principle capable of resolving that drift whilst preserving human judgement. When combined with the existing structure of hybrid intelligence and governance frameworks, a specific distribution of decision authority follows. What the field needed was not another account of how human and AI should work in concert, but a clear answer to who decides. What the literature has described as interaction, collaboration, and involvement is here resolved into authority.

4. The Hybrid Decision-Making Model

Section 3 established three findings in sequence: that the absence of decision authority produces predictable drift, that augmentation is the only principle capable of resolving that drift, and that applying that principle to the existing hybrid intelligence and governance frameworks produces a specific distribution of decision authority. This section presents that distribution. It is the Hybrid Decision-Making Model.

The HDM is not a set of options from which institutions select according to preference. It is a structure that locates every decision in AI-integrated higher education practice within one of three defined tiers. The tiers do not classify activities or workflows. They assign authority: who holds the right to make the final judgement, and on what basis. A decision is not negotiated into a tier. It belongs there because of what it is. Tier assignment is a structural determination based on the nature of the decision. Efficiency is a consequence of correct assignment, not a basis for it. An institution that assigns a decision to a tier because it is convenient to do so has not applied the HDM. It has misapplied it.

Tier assignment is determined by two criteria from Section 3.3: the consequence attached to the decision and the contextual and relational knowledge it requires. Both criteria are stable. What changes across institutional contexts is not the criteria but their application. A decision that is routine in one context may carry higher consequence in another and will be assigned to a higher tier accordingly. That is principled application of stable criteria, not situational flexibility. As AI capability and institutional context evolve, tier assignments should be reviewed. The criteria for assignment do not change.

The three tiers are related in three distinct ways. They are hierarchical in terms of consequence and human authority: Tier 3 carries more consequence and requires more human authority than Tier 1. They are complementary in terms of function: each tier handles a category of decisions that the other two cannot handle well, and collectively the three tiers cover the full decision environment of AI-integrated higher education practice. They are simultaneous in terms of operation: a teaching day does not proceed through Tier 1 decisions, then Tier 2, then Tier 3 in sequence. All three tiers operate concurrently across the full range of decisions an educator faces.

The pyramid structure (Figure 2) reflects this logic directly and its shape is not arbitrary. The base is wide because Tier 1 decisions are numerous: in any AI-integrated educational environment, the majority of decisions that AI can assist with are routine, high-volume, and low-consequence. The middle band is narrower because Tier 2 decisions require the specific condition of necessary integration, which applies to a smaller range of decisions than routine processing but a larger range than those carrying irreversible ethical consequence. The apex is narrow because Tier 3 decisions are the fewest in number and the highest in consequence: their specificity, ethical weight, and relational complexity mean that they cannot be routinised, scaled, or delegated.

Before the three subsections begin, a practical tier assignment test follows from the two criteria directly. When a decision must be placed, three questions apply in sequence. First: if an error in this decision is not corrected, can the student recover from its consequences without lasting harm, and does the decision call for no interpretive judgement that only the educator can provide? If yes: Tier 1. Second: does this decision require both AI-provided data and the educator's contextual knowledge to be formed well, such that neither contribution alone produces a valid outcome? If yes: Tier 2. Third: does the validity of this decision depend on human judgement in a way that cannot be delegated, because its consequences are irreversible, its ethical weight is significant, or its relational specificity exceeds what data can capture? If yes: Tier 3. When a decision sits at the boundary between two tiers, assigning too low risks structural drift. The HDM is designed to prevent that. The examples across the three subsections are drawn primarily from assessment-adjacent contexts for illustrative clarity. The HDM applies across the full range of educational decision-making in AI-integrated practice.

The three tiers are presented in full in the subsections that follow. Figure 2 provides a structural overview before the subsections begin.

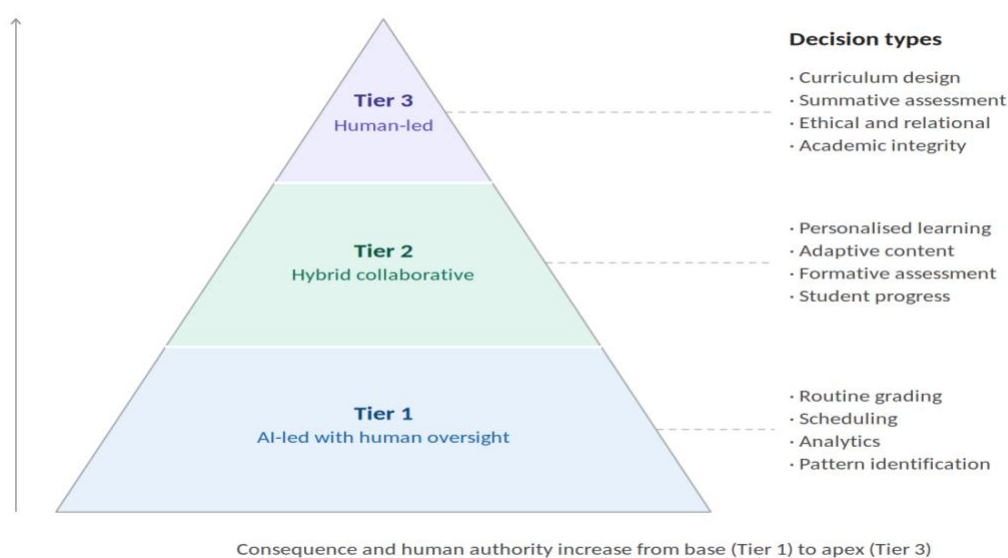


Figure 2: The Hybrid Decision-Making Model: distribution of decision authority across three tiers of AI-integrated higher education practice.

Tier 1 (base) covers routine, reversible decisions where AI leads and the human retains oversight. Tier 2 (middle) covers decisions requiring both data and contextual judgement, where the human holds final authority. Tier 3 (apex) covers decisions of high consequence and ethical weight, where human primacy is a structural requirement. Decision consequence and human authority increase from base to apex. Decision types listed are illustrative, not exhaustive.

4.1. Tier 1: AI-Led with Human Oversight

Tier 1 covers decisions that are routine, repeatable, and low in consequence. Human oversight at this tier does not mean passive monitoring. It means retained authority: the educator's unconditional right to review, question, and overrule any AI output.

Low consequence, as used here, refers to the reversibility of the outcome: where an error is recoverable without lasting harm to the student, and where correcting it calls for no interpretive judgement that only the educator can provide, the decision falls within Tier 1. A decision is not Tier 1 because it is convenient to automate. It is Tier 1 because its consequence is genuinely recoverable and its assessment requires no contextual or relational knowledge that only the educator holds. Where decisions of this nature recur at high volume and follow predictable patterns, AI is better positioned to execute them efficiently and consistently. Efficiency and consistency are the appropriate standards here, because these are decisions that should not vary according to who performs them. Routine grading of objective assessments, scheduling, analytics, and pattern identification share these properties. They are listed as illustrations of a category, not as a complete inventory. This configuration is not a design preference. It is the direct expression of the human-over-the-loop arrangement identified by IMDA and PDPC (2020), applied to the category of decisions where consequence is lowest and AI capability is highest.

The governing principle of this tier is clear: AI performs the task, but the human retains the right to decide whether the outcome stands. That is what separates Tier 1 from AI replacement. The task has been delegated. The authority has not. What keeps authority with the human is structural assignment, not personal vigilance. Because the decision has been categorised as Tier 1 in advance, the structure signals that every AI output is subject to human review.

The frequency with which that authority is exercised does not determine where it lies. An educator who reviews one in twenty AI-generated outputs has not transferred authority for the other nineteen. Authority is held structurally. Its exercise is discretionary. Its presence is unconditional. Most educators working with AI tools will

recognise this intuitively. The question is not whether they trust the output. It is whether they know they can change it. The HDM does not specify a minimum review threshold, because doing so would turn a decision authority framework into a compliance checklist. The model establishes where authority resides. Institutional practice determines how it is used.

This arrangement addresses the meta-decision burden identified in Section 3.1. At Tier 1, the question of whose decision this is has already been answered by the structure. The educator receives AI output already framed as subject to review. That framing removes the meta-decision, and in doing so interrupts the first and most pervasive condition under which drift takes hold.

Tier 1 expresses the augmentation principle established in Section 3.2 as efficiency with oversight. AI leading routine, high-volume tasks does not displace the educator. It frees attention for the decisions in Tier 2 and Tier 3 that require interpretation, ethical judgement, and relational knowledge that AI cannot provide. The HDM does not prevent educators from engaging closely with Tier 1 outputs. An educator who reviews AI-generated scores in detail retains full access to the data they contain. The base of the pyramid (Figure 2) is wide because Tier 1 decisions are numerous, and managing them through AI leadership with human authority retained is what makes the higher tiers of the model workable in practice. The structural safeguard described here operates at the level of individual decision-making. Whether institutions maintain the conditions under which that safeguard remains effective over time is a question addressed in Section 6.

4.2. Tier 2: Hybrid Collaborative

Tier 2 is the most conceptually demanding tier in the HDM. Where Tier 1 assigns leadership to AI and Tier 3 assigns leadership to the human, Tier 2 requires both. At Tier 2, collaboration does not mean that both human and AI are involved. It means that the decision cannot be resolved without both. This is what the HDM means by collaboration at Tier 2: not joint presence but necessary integration. The decision requires both contributions, and neither alone is sufficient to produce a valid outcome.

The decisions that fall within Tier 2 share a specific structural condition: data without context is incomplete, and context without data is insufficient. AI identifies patterns across student performance, engagement, and progress at a scale that lies beyond the capacity of individual human observation. The educator holds contextual and relational knowledge about each student that cannot be inferred from data alone. Without AI input, the educator is deciding without the full information the decision requires. Without human input, the AI output cannot stand as a decision at all, because it lacks the interpretive and ethical authority that makes a decision valid in an educational context. These are not equivalent absences. One leaves the decision incomplete. The other leaves it without authority. The second absence is a structural claim, not a technological one: within the HDM, a decision whose validity requires human interpretive authority cannot be valid without it, regardless of how advanced AI becomes. Personalised learning, adaptive content, formative assessment interpreted across time, and student progress all share this condition. This configuration corresponds to the human-in-the-loop arrangement identified by IMDA and PDPC (2020), applied to the category of decisions where both data and contextual knowledge are constitutive of the outcome.

Authority at Tier 2 is pre-assigned to the human, as it is across all three tiers of the HDM. What the collaboration produces is not authority but the information base on which that pre-assigned authority can be fully and responsibly exercised. The human cannot exercise that authority well without AI input. The AI cannot hold that authority at all. Final judgement is exercised at the point where data and context must be integrated into a specific decision. Before that moment, both contributions are active. At that moment, the human decides. The collaboration is therefore structured and asymmetric: both contributions are necessary in process, but authority is unambiguously human at the point of decision. Human authority is final at Tier 2 not only because the HDM assigns it there, but because the decisions at this tier carry moral weight that requires a locus of accountability that only a human professional can provide.

What separates Tier 2 from Tier 1 is not the degree of human involvement but its nature. At Tier 1 the human oversees. At Tier 2 the human integrates. AI contributes patterns that are not visible at the level of individual observation. The educator contributes contextual understanding that cannot be inferred from data alone. The decision emerges from the integration of both and cannot be reduced to either. AI identifies a pattern across a student's formative assessment results. The educator interprets what that pattern means for this student, given what is known about how that student thinks, engages, and responds. The AI produced the pattern. The educator produced the decision. Remove the pattern and the decision lacks its evidential basis. Remove the educator and the pattern has no one to interpret it and no authority to act on it. Every educator who has ever looked at a spreadsheet of student scores and thought: but that is not the whole story, already understands what Tier 2 is for.

Two collapse conditions define the boundaries of this tier and serve as its diagnostic tests. If AI dominates and the human only reviews the output, the decision has migrated to Tier 1. If the human proceeds without AI input, the decision has migrated to Tier 3. Neither migration is a failure of the model. Both are visible because the model has defined what Tier 2 requires. The HDM makes migration identifiable. What separates Tier 2 from Tier 3 is not the weight of the decision but the role of AI input within it. At Tier 2 the decision cannot be formed without AI. At Tier 3 it can, and must, be made without AI determining its direction.

Tier 2 also addresses the cognitive dimension identified in Section 3.1. Without AI handling the data-processing burden, the cost of producing a fully informed decision would exceed what most educators can sustain across a working day. AI at Tier 2 reduces that cost whilst the structure preserves the educator's authority over what the information means. This is the augmentation principle from Section 3.2 operating as informed judgement: the educator decides, and the AI ensures the decision is made with the fullest possible evidential foundation.

4.3. Tier 3: Human-Led with AI Support

If Tier 1 is defined by efficiency and Tier 2 by integration, Tier 3 is defined by responsibility. Responsibility, at Tier 3, refers to the irreversible, ethically consequential, and relationally specific nature of the decisions this tier contains. These three characteristics are the forms that high consequence and deep contextual knowledge take at their most demanding expression. Curriculum design, summative assessment, ethical and relational judgement, and academic integrity belong here not because AI is incapable of contributing to these areas, but because the validity of these decisions depends on human authority. That dependency is not a policy preference. It is a structural consequence of what these decisions are. This configuration corresponds to the human-led arrangement identified across the theoretical foundations of Section 3, where the nature of the decision makes human primacy not a structural preference but a structural requirement.

The distinction between Tier 2 and Tier 3 must be stated directly. At Tier 2, AI contributes to forming the decision: the human cannot decide well without the AI's contribution, and the decision emerges from their integration. At Tier 3, AI contributes to informing the human who decides: the AI provides input, but the decision belongs to the human regardless of what that input contains. At Tier 2, AI input is constitutive. At Tier 3, it is informational. Remove AI from a Tier 2 decision and the decision is incomplete. Remove AI from a Tier 3 decision and the decision is still valid, because its validity derives from human judgement, not from data. AI improving a Tier 3 decision does not make the input constitutive. Augmentation improves the judgement. Integration requires it. The first is Tier 3. The second is Tier 2.

AI is present at Tier 3. It is not excluded. What is bounded is its role. The HDM does not claim that AI information has no effect on the educator's thinking at Tier 3. Information always enters a frame of judgement and affects it to some degree. The claim is more precise: influence without authority is not a structural problem for the HDM, because the model does not require AI to have no effect. It requires that the effect itself does not constitute authority. An educator influenced by AI-provided data has not had their authority compromised. The structure determines the locus of authority, not the source of influence. At Tier 3, the authority to determine how AI-provided information is weighted, contextualised, and acted upon rests entirely and unconditionally with the human. AI provides the input. The educator owns the interpretation.

The boundary at Tier 3 is enforced structurally, not personally. Because the decision has been identified as Tier 3 in advance, the structure signals that AI output is informational only. The educator does not need to resist AI influence in the moment. The frame has already been set by the structure. This connects directly to the meta-decision removal established in Section 3.1: the educator does not need to ask whether this is their decision to make. Tier 3 resolves the most consequential form of drift: the uncritical acceptance of AI outputs at the moment of highest stakes.

The academic integrity example deserves specific attention because it is the decision type most likely to generate a reviewer challenge. AI tools that detect patterns associated with academic integrity concerns are performing a Tier 2 function. They are contributing data to the information base. The decision about whether a breach has occurred, what it means for this student, and what consequence is appropriate is a Tier 3 decision. The two steps happen in sequence. The first is Tier 2. The second is Tier 3. In practice, institutions often treat them as a single process. The HDM's value is precisely that it separates them. That separation is not a theoretical refinement. It is a structural protection for the student.

Tier 3 expresses the augmentation principle from Section 3.2 at its most essential. The educator is freed to do what only the educator can do: exercise the irreversible, ethically grounded, relationally informed judgement that gives these decisions their validity. The base of the pyramid is defined by volume. The middle is defined by integration. The apex is defined by responsibility. The HDM does not ask institutions to choose between efficiency and human judgement. It asks them to recognise that some decisions derive their legitimacy entirely from the fact that a human made them, and to protect that legitimacy by design.

The three tiers collectively constitute a complete structure for the distribution of decision authority across AI-integrated higher education practice. They cover the full decision environment, because the two criteria that govern tier assignment apply to every decision an AI-integrated educational practice contains, and the three tiers are the complete and non-redundant set of configurations those criteria yield.

The HDM's scope within this paper is classroom-level decision-making in AI-integrated teaching practice. Decisions at the programme level, departmental level, or institutional governance level involve different actors, different accountability structures, and different consequence profiles. The HDM does not extend to those levels here. That boundary is a deliberate scope decision that future research may extend.

Educational workflows frequently contain decisions across multiple tiers within a single process. A student progress review may begin as a Tier 2 data interpretation and end as a Tier 3 welfare judgement. The HDM assigns authority at the level of each individual decision within the process, not at the level of the process as a whole.

Two forms of movement between tiers must be distinguished. Contextual reassignment is expected and principled: the same category of decision may be assigned to a higher tier in one institutional context than another because the consequence dimension has changed. Practical drift is problematic: a decision correctly assigned to Tier 2 begins to be treated as Tier 1 because AI outputs are accepted without genuine integration. The HDM is designed to prevent the second form of movement by removing the meta-decision burden.

The augmentation principle operates specifically and differently at each tier. At Tier 1 it is efficiency with oversight. At Tier 2 it is informed judgement. At Tier 3 it is freed authority. These three expressions show that augmentation is not applied once to the model as a whole. It operates at each tier according to the nature of the decisions that tier governs.

The HDM is not one possible response to the conditions Section 3 identified. It is the response those conditions require: the structure that removes the meta-decision burden bounded rationality produces, operationalises the augmentation principle across the full range of educational decisions, and derives its three tiers from the only stable configurations that the interaction of consequence and human involvement yields. Given the premises established in Sections 2 and 3, the HDM is not a useful addition to the field. It is what the field's own foundations make necessary.

5. Empirical Validation

Sections 3 and 4 established the HDM through theoretical derivation and structural presentation. This section asks a different question: does the model reflect what educators were already doing before it existed? The study that provides the answer is Chin (2025), a qualitative investigation conducted through semi-structured interviews with seven experienced educators in a Southeast Asian higher education context. The study was conducted in the absence of formal institutional guidance on AI integration, which is precisely the condition the HDM is designed to address. Nine themes emerged across three categories: pedagogical implications, operational considerations, and ethical and professional dimensions, through a thematic analysis following Braun and Clarke's (2006) six-step inductive procedure. The educators who generated those themes had no knowledge of the HDM. They were describing their own practice, drawing their own boundaries, and making their own distinctions about which decisions AI should lead, which should remain theirs, and which required both.

What the mapping in this section demonstrates is not that the HDM fits the data. It is that the data reflects the HDM: the three tiers were already present in educator practice before the model existed to name them. That is not a coincidence of design. It is the strongest form of confirmation available. Figure 3 presents the full mapping before the subsections examine each tier in turn.

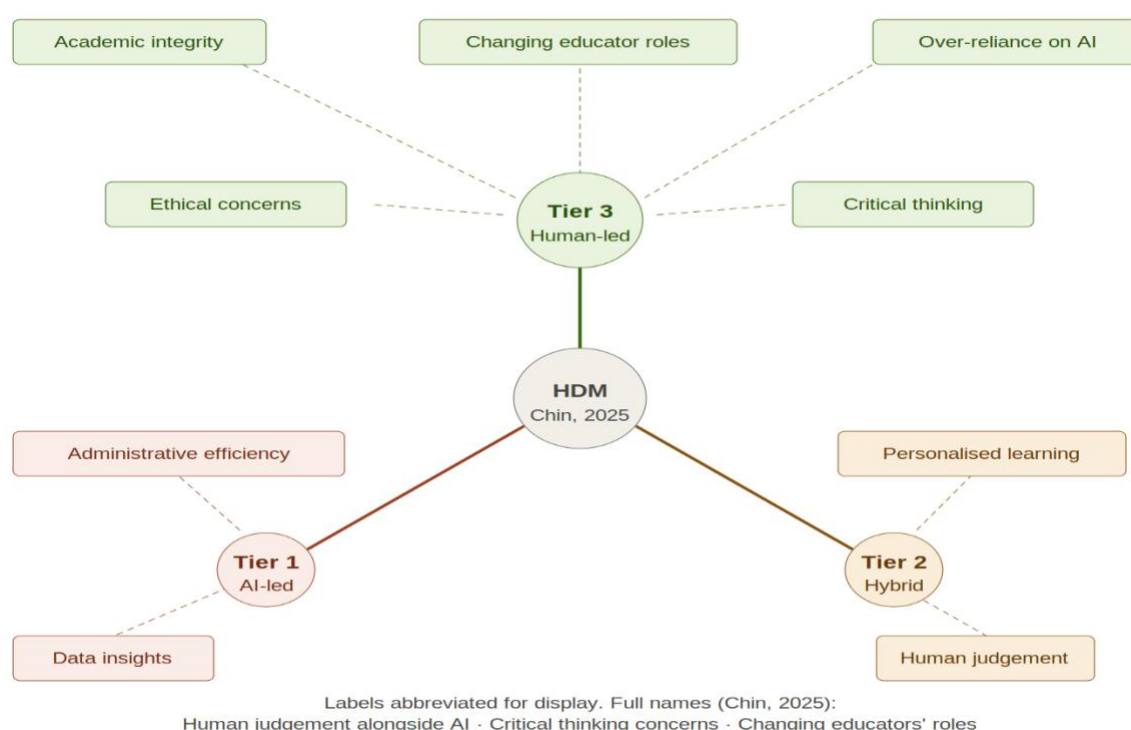


Figure 3: Hub-and-spoke diagram mapping nine empirical themes from Chin (2025) to the three tiers of the Hybrid Decision-Making Model.

Coral indicates Tier 1, amber Tier 2, and green Tier 3, with two themes mapped to Tier 1, two to Tier 2, and five to Tier 3. Abbreviated labels include Human judgement alongside AI, Critical thinking concerns, and Changing educators' roles. The mapping shows that themes from all three original categories of the study align with the HDM's tiers, indicating that the model's structure reflects patterns of practice across the dataset.

5.1. Tier 1 Validated: Administrative Efficiency and Data Insights

The two themes that map to Tier 1 are administrative efficiency and data insights, both drawn from the operational considerations category of Chin (2025). Their placement at Tier 1 rests on a structural correspondence: both themes describe decisions that are routine, repeatable, pattern-based, and low in consequence, which are precisely the conditions that define Tier 1 in the HDM.

The educators were not simply accepting AI outputs uncritically, nor were they redoing the work themselves. They were retaining the authority to review, question, and overrule whilst allowing AI to execute. Participant 1 described this explicitly: AI manages routine procedural elements whilst human expertise guides evaluation strategies. That is Tier 1 in practice. The task is delegated. The authority is not.

This pattern confirms what Tier 1 formalises. Educators had already separated execution from authority in routine decisions before any formal structure existed to require it.

Educators were consistently directing their cognitive attention towards the decisions that required it most, whilst allowing AI to handle the high-volume, pattern-based tasks that did not. That is bounded rationality working in the educator's favour: not as a source of drift, but as a natural allocation of cognitive resource that the HDM now makes structurally explicit and institutionally replicable.

5.2. Tier 2 Validated: Personalised Learning and Human Judgement Alongside AI

Two themes from Chin (2025) correspond to Tier 2: personalised learning and human judgement alongside AI. The governing claim is clear. The educators were not simply using AI to support their decisions. They were relying on AI to extend what they could know, whilst retaining full responsibility for what that knowledge meant for each student. AI revealed the pattern. The teacher decided what that pattern meant for the student. That is Tier 2, and it was already present in practice before the model existed to name it.

Personalised learning was not a general preference for individualised instruction. It was a specific practice in which AI-generated data about student performance and progress was combined with the educator's knowledge of each student as a person. AI sees the pattern across the data. The educator knows what that pattern means for this particular student. Remove the pattern and the educator is working without the information needed to decide well. Remove the educator and the pattern has no one to interpret it and no authority to act on it. This is why personalised learning belongs at Tier 2. It was placed there because the structural condition matches.

Human judgement alongside AI makes the same point through the educators' own words. What the interviews described was a practice in which AI extended what educators could know, whilst they retained responsibility for what those insights meant for each student. Participant 3 described how shifting routine tasks to AI created space for teaching decisions that depended heavily on expert knowledge and an understanding of how each student was progressing. That is not AI supporting a decision the educator could have made without it. That is AI making a form of decision possible that would otherwise be incomplete. The collaboration was not optional. It was constitutive.

A teacher who has sat with a data report at the end of a long week, trying to work out which student needs what attention next, already knows this. The data narrows the field. The knowledge of the student determines the response. Neither alone is enough.

They arrived at this arrangement through their own professional judgement about what good practice required. The HDM does not prescribe a new arrangement. It names and formalises a mode of practice that was already operating consistently in the data.

5.3. Tier 3 Validated: Academic Integrity, Ethical Concerns, Critical Thinking Concerns, Over-Reliance on AI, and Changing Educators' Roles

Five themes from Chin (2025) correspond to Tier 3, and what unites them is not their content but their consequence. Every one of them involves a moment where the educator cannot hand the decision to AI, because the decision is about a person and the educator is the one who will be held responsible for it.

The five themes are academic integrity, ethical concerns, critical thinking concerns, over-reliance on AI, and changing educators' roles. Each is a judgement about a different dimension of practice: fairness and consequence, right and wrong, student development, dependency and autonomy, and professional responsibility. But the structural condition is the same across all five. The decision cannot be reduced to data, its consequence cannot be reversed, and its validity depends on the fact that a human made it. The AI may provide a similarity score. The decision about what that score means for this student, whether a breach has occurred, what consequence is appropriate, and what the student needs, is not produced by the score. It is produced by the educator's judgement about a person, a situation, and what fairness requires. These are the conditions that Section 3.3 placed at Tier 3: irreversible consequences, ethical weight, and relational specificity that exceeds what data can capture.

The interviews in Chin (2025) show that educators were already acting as though these decisions were theirs to make. They questioned AI outputs where the consequence for the student was significant. They retained personal responsibility for judgements about academic integrity rather than treating detection tools as definitive. They described protecting students' capacity for independent thinking as a professional obligation that AI efficiency could not replace. Nobody told them to draw this boundary. They drew it because the nature of the decisions required it.

The consistency with which these five themes cluster around decisions of consequence is itself significant. The educators had no formal structure to follow. Yet they were consistently distinguishing between what AI can indicate and what only a human can decide. The boundary between AI input and human authority was already being drawn in practice before the HDM existed to name it. The earlier subsections showed that educators had already separated execution from authority and integrated data with contextual judgement. This subsection shows that they already knew which decisions were theirs alone to make. The model does not impose that boundary. It reveals the one educators were already defending.

The educators in Chin (2025) were not applying the HDM. They were responding to the structural conditions of their own practice: the reversibility of some decisions, the constitutive role of data in others, and the irreversible ethical weight of a third category. Those conditions organised their behaviour before the model existed to name them. The HDM does not describe educator behaviour. It reveals the structure that was already organising it. If a reviewer asks whether the educators knew they were operating within these tiers, the answer is no. But their decisions show that they were.

6. Reflections on Institutional Practice

Sections 3 and 4 established the HDM as a necessary structure. What remains is a different question: not whether the structure is valid, but what must be true for it to remain so. Section 3.1 noted that structural safeguards operate at the level of individual decisions, and that whether institutions maintain the conditions for those safeguards to hold over time is a question the model identifies but does not resolve within Section 4. This section addresses that question by naming the conditions the model's own logic requires.

6.1 Sustainability and the Institutional Condition

The HDM resolves drift at the level of individual decisions by removing the meta-decision burden. That resolution holds because the tier assignment is already in place before the decision arises. The same logic applies at the institutional level, and the consequences of its absence are more serious because they operate at scale.

Institutional drift accumulates through repeated decision cycles in which authority is left unclear, boundaries remain informal, and responsibility stays implicit. What accumulates is a quiet realignment of where educational judgement is exercised, away from deliberate human authority and towards whatever the AI has already produced. Decisions that belong at Tier 2 begin cycling as Tier 1. Decisions that belong at Tier 3 begin cycling as Tier 2. The structure has not broken. It has drifted.

The institutional equivalent of removing the meta-decision is the formal assignment of decision authority as part of governance, established before practice begins. Training programmes and policy statements address what educators know about the problem. They do not change the conditions that produce it. For the HDM's structural logic to hold, three conditions must be true. Decision authority must be formally assigned to tiers before educators encounter those decisions. Those assignments must be visible at the point of practice. And they must be reviewed as AI capability and context evolve. Without these conditions, the responsibility for determining decision authority returns to the individual educator, and the conditions for drift are reinstated at scale. What the HDM does for the individual, the institution must do for the system.

6.2 Responsibility and the Maintenance of Structure

The HDM transfers the responsibility for maintaining decision authority assignments to the institution. That transfer does not occur automatically. It requires institutional responsibility for monitoring whether the structure is being followed, identifying when practice has drifted, and reviewing assignments when conditions change. This is not supervision of individual decisions. It is maintenance of the structure that governs them, a governance function that operates above the level of individual practice.

A structure that is not actively maintained will not hold. Interpretation shifts. Practice changes. AI capability evolves. Without a clear structure for determining whose responsibility it is to maintain the assignment, that determination defaults to those closest to the decision at the moment it arises.

Singapore's Model AI Governance Framework (IMDA and PDPC, 2020) and Brunei Darussalam's national guide on AI governance and ethics (AITI, 2025) both treat the determination of human involvement in AI-augmented decision-making as a core organisational responsibility. Research examining how AI governance frameworks translate into policy practice confirms that transparency and accountability in decision-making remain central concerns across national contexts (Yar et al., 2024). The HDM specifies what that involvement looks like at the level of classroom practice. What those frameworks call for at the organisational level, the HDM provides at the practice level, requiring in return that institutions treat tier maintenance as a governance responsibility rather than an individual obligation.

Without a clear locus of responsibility for maintaining the structure, the assignment of decision authority gradually becomes interpretive again. The HDM removes uncertainty from decisions. Institutions must remove uncertainty from the structure that governs them over time.

6.3 Stability, Review, and the Protection of Principles

The HDM's criteria for tier assignment are stable. A decision belongs at Tier 3 not because current AI cannot handle it, but because its irreversible consequence and relational specificity make human authority structurally necessary. That structural necessity is independent of AI capability. The principles remain constant. What changes is how they apply.

As AI capability evolves and institutional contexts shift, decisions may require reclassification. That reclassification is a feature of principled application. What is a problem is when reclassification occurs informally, without being tested against the two criteria. That is not adaptation. It is unexamined reassignment of authority, and it reproduces the conditions for drift.

The HDM therefore requires deliberate review as a structural condition of its own validity. Principled reassignment is deliberate and tested. Informal reassignment is drift by another name.

Change is inevitable. Unexamined change is drift. The HDM remains valid not because it adapts to change, but because it requires its application to be deliberately reviewed so that its principles continue to govern under changing conditions. Structure must exist. Responsibility must maintain it. Review must protect it from silent change.

Section 6 has established three conditions the HDM's own logic requires. The structure of decision authority must be assigned before practice begins. That assignment must be maintained as an institutional responsibility. And its application must be deliberately reviewed as conditions change. These are not recommendations. They are the conditions without which the model's structural logic does not survive contact with institutional reality. The HDM does the work of clarifying whose decision it is. What keeps that clarity intact over time is a different order of work entirely, and it belongs to the institution.

7. Discussion

This paper argues that the central problem in AI-integrated education is not technological but structural: not what AI can do, but who decides. When AI contributes to educational decisions, authority over those decisions shifts, and that shift has largely occurred without being named or governed. Epistemic authority, as used here, refers to the recognised right to determine what counts as valid judgement in an educational setting: whose assessment stands, whose interpretation holds, and whose decision defines the outcome for the student. What the field has treated as a question of capability and risk, this paper reframes as a question of authority.

Three bodies of scholarship established that human and AI must collaborate deliberately, that the teacher must remain central, and that task allocation must be designed rather than assumed. Each stopped at the same point: they specified how work should be shared but did not specify who holds authority when a judgement must be made. The HDM begins where that stops. It provides what the existing literature established as a necessary condition but did not supply: a structure for assigning decision authority explicitly. In doing so, it completes a line of work the field had already begun, even if the field had not yet named what it was building towards.

This contribution is not only theoretical. By naming decision authority as the structural question at the centre of AI integration, it connects classroom practice directly to the governance conversations in which accountability depends on clearly defined human roles. Independent work confirms that this connection is timely: a scoping review of 32 empirical studies found that teachers were already being asked to determine what AI can and cannot do in assessment, without a structure to support that determination (Xia et al., 2024). The HDM provides that structure. For practitioners, it reframes AI use not as a matter of preference or efficiency, but as a matter of structured judgement and responsibility. The field can now move from recognising that AI affects decisions to specifying how those decisions should be governed. It can also evaluate AI use not only in terms of effectiveness or ethics, but in terms of whether decision authority is appropriately assigned.

The HDM specifies a conceptual structure for decision authority but does not prescribe mechanisms for institutional implementation. It is developed within a higher education context and may require adaptation when applied to other educational levels or sectors. Its empirical grounding in Chin (2025) confirms alignment in practice but does not establish generalisability across institutions or national settings. The framework assumes evolving AI capability but does not empirically test how technological change alters tier assignment over time. These boundaries do not diminish the contribution. They define what it sets out to do, and what it leaves, deliberately, for the work that follows.

This paper has not added another framework to a field that already has many. It has made visible the structure that was missing from all of them.

8. Future Research

The HDM makes decision authority visible. The most immediate research priority is the development of an instrument through which decision authority can be observed and measured in AI-integrated educational practice: one that allows researchers to identify which tier a decision belongs to, whether the correct authority is being exercised, and where drift is occurring without being named. Without that instrument, the questions that follow cannot be pursued empirically. What becomes possible to study now that decision authority is visible is the question this section addresses.

The first set of directions faces outward. How do institutions translate the HDM's tier assignments into governance practice? Who holds responsibility for maintaining them, and how do national frameworks such as AITI (2025) and IMDA and PDPC (2020) connect to classroom-level decision authority? These questions concern implementation and sustainability, the conditions Section 6 identified but did not resolve. They also extend to cross-national validation: whether educators in other higher education systems, working without formal guidance, draw the same structural distinctions the HDM formalises remains to be tested. And as AI capability evolves, how do institutions review tier assignments deliberately rather than allowing informal drift to substitute for principled reassignment? These are the questions closest to practice, and they are the ones the field is most immediately positioned to pursue.

The second set of directions faces inward. The HDM's three-tier structure gives researchers an instrument for observing how decision authority is actually distributed in AI-integrated classrooms, where it resides, where it has drifted, and when AI input moves from informational to constitutive without that movement being named. That observational capacity is new. Equally important is the question of whether educators feel safe exercising the authority the structure assigns. The HDM specifies that Tier 3 authority belongs to the educator unconditionally. But holding authority and feeling safe to use it are not the same condition. Whether educators feel safe challenging an AI-generated output, particularly in institutional environments where that output carries implicit endorsement, is a question the HDM makes possible to ask but does not resolve. The relationship between structural authority assignment and the psychological conditions under which that authority is exercised, explored in the context of educational leadership by Edmondson (1999) and Chin (2026), represents a necessary and largely unexamined direction for future empirical work. A structure can assign authority. It cannot assign the willingness to bear its weight. An educator who holds Tier 3 authority but does not feel safe challenging an AI-generated output has not lost the authority. The authority simply has not been used. Whether that happens, and why, is a question the HDM makes possible to ask and necessary to answer.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest: The author declares no conflicts of interest with respect to the research, authorship, and/or publication of this article.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F., van Hoof, H., van Riemsdijk, B., van Wynsberghe, A., Verbrugge, R., Verheij, B., Vossen, P., & Welling, M. (2020). A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(8), 18–28. <https://doi.org/10.1109/MC.2020.2996587>
- Authority for Info-communications Technology Industry of Brunei Darussalam. (2025). *Guide on artificial intelligence (AI) governance and ethics for Brunei Darussalam*. AITI. <https://www.aiti.gov.bn/guide-books/guide-on-ai-governance-and-ethics-for-brunei-darussalam/>
- Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, 26(2), 600–614. <https://doi.org/10.1007/s40593-016-0105-0>
- Borchers, C., Viberg, O., & Kizilcec, R. F. (2026). Who decides in AI-mediated learning? The Agency Allocation Framework. In Proceedings of the Thirteenth ACM Conference on Learning @ Scale (L@S '26). ACM. <https://doi.org/10.1145/3774398.3811627>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

- Bredeweg, B., & Kragten, M. (2022). Requirements and challenges for hybrid intelligence: A case-study in education. *Frontiers in Artificial Intelligence*, 5, 891630. <https://doi.org/10.3389/frai.2022.891630>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20, 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Chin, P. P. L. (2025). AI in Higher Education: A Case Study Examining Decision-Making Shifts and Personalised Learning Through Educators' Perspectives. *Education Quarterly Reviews*, 8(2), 23–39. <https://doi.org/10.31014/aior.1993.08.02.576>
- Chin, P. P. L. (2026). The Pause Before the Answer: Psychological Safety as the Missing Condition Between Visionary Leadership and Genuine Teacher Voice. *Education Quarterly Reviews*, 9(2), 1–10. <https://doi.org/10.31014/aior.1993.09.02.712>
- Chou, C. Y., Huang, B. H., & Lin, C. J. (2011). Complementary machine intelligence and human intelligence in virtual teaching assistant for tutoring program tracing. *Computers and Education*, 57(4), 2303–2312. <https://doi.org/10.1016/j.compedu.2011.06.005>
- Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. <https://doi.org/10.2307/2666999>
- Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework*. Stanford Research Institute. <https://www.doungelbart.org/pubs/augment-3906.html>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Holstein, K., & Aleven, V. (2022). Designing for human–AI complementarity in K–12 education. *AI Magazine*, 43, 239–248. <https://doi.org/10.1002/aaai.12058>
- Holstein, K., Aleven, V., & Rummel, N. (2020). A conceptual framework for human-AI hybrid adaptivity in education. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial intelligence in education*, 240–254. Springer. https://doi.org/10.1007/978-3-030-52237-7_20
- Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher-AI complementarity. *Journal of Learning Analytics*, 6(2), 27–52. <https://doi.org/10.18608/jla.2019.62.3>
- Infocomm Media Development Authority and Personal Data Protection Commission. (2020). Model AI governance framework (2nd ed.). <https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/factsheets/2020/model-ai-governance-framework>
- Miao, F., & Holmes, W. (2021). *Artificial intelligence and education: Guidance for policymakers*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- Paiva, R., & Bittencourt, I. I. (2020). Helping teachers help their students: A human-AI hybrid approach. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial intelligence in education*, 448–459. Springer. https://doi.org/10.1007/978-3-030-52237-7_36
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12, 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Schiff, D. (2022). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education*, 32(3), 527–563. <https://doi.org/10.1007/s40593-021-00270-2>
- Simon, H. A. (1947). *Administrative behavior*. Macmillan.
- Simon, H. A. (1990). Bounded rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Utility and probability*, 15–18. Macmillan.
- Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21, 40. <https://doi.org/10.1186/s41239-024-00468-z>
- Yar, M. A., Hamdan, M., Anshari, M., Fitriyani, N. L., & Syafrudin, M. (2024). Governing with intelligence: The impact of artificial intelligence on policy development. *Information*, 15(9), 556. <https://doi.org/10.3390/info15090556>

Investigating Chatbot-Assisted Self-Regulated Learning of Basic Electrical Engineering: Effectiveness and Variations of Students' Learning Outcomes

Hamonangan Tambunan¹, Efendi Napitupulu², Wanapri Pangaribuan³

^{1,2,3} Universitas Negeri Medan, North Sumatra, Indonesia

Correspondence: Hamonangan Tambunan, email: hamonangantambunan@unimed.ac.id
ORCID ID: <https://orcid.org/0000-0002-7635-2112>

Abstract

This study examines the effectiveness of a learning model for prospective electrical engineering teachers, who often encounter obstacles stemming from differences in school types and statuses. Their initial abilities do not support learning, so we develop a learning model to accommodate them and investigate their learning outcomes across groups by school origin and gender. The research design uses quantitative analysis of learning outcomes and chatbot usage, along with qualitative student feedback, in a mixed-methods approach. The study involved 159 students (6 classes) based on their school origin: school status (96 public, 63 private) and school type (106 senior high school, 53 vocational high school). There were 108 males and 51 females. We collected data through tests and questionnaires and analyzed it using ANOVA in IBM SPSS v26. We found no difference in students' learning outcomes by origin school type and status, but females performed better than males. It is shown that Chatbot-assisted self-regulated learning is practical.

Keywords: Electrical Education, Learning Outcome, Self-Regulated Learning

1. Introduction

Educational chatbots are widely accepted tools for boosting student involvement in virtual learning environments. The research has received substantial attention because self-regulated learning can be developed among students through these methods (Guan et al., 2024). These conversational agents offer personalized, interactive experiences that can be particularly beneficial. In technical disciplines, basic electrical engineering often requires students to make multiple attempts to grasp complex concepts. The student will receive feedback that addresses their individual needs and aligns with their current level of understanding (Chang et al., 2023). AI-driven chatbots can perform the following functions. The implementation of Artificial Intelligence (AI) in educational settings leads to substantial improvements in student learning outcomes, engagement, and interest levels. The program can assist students interested in studying Science, Technology, Engineering, and Mathematics. Its ability to provide instant

feedback and individualized guidance allows learners to take an active role in their educational path, developing a stronger connection to their learning process and a deeper understanding of the subject matter (Yin et al., 2025). A personalized approach enables students to learn essential self-regulation techniques—goal setting, self-assessment, and metacognitive monitoring—that are vital to academic success. Lan and Zhou (2025) examine the factors that lead to success in higher education, stating that large language Model-based chatbots demonstrate substantial capability to improve various teaching methodologies, including flipped classrooms. The teacher can maintain active communication with students and promptly answer the students' questions (Vera, 2024). Nevertheless, technological progress can create various obstacles for society to overcome, and opportunities for educators, necessitating a re-evaluation of pedagogy. It is crucial to examine how effective the chatbot implementation is in driving students' self-regulated learning, especially in basic electrical engineering, across schools of varying origins.

So far, there has been little research on students' learning outcomes across different types of schools and statuses when engaging in similar learning models. Electrical engineering education students in Indonesia come from various schools and backgrounds, possess different initial competencies, and have unique prior knowledge and learning preferences. Recognizing and understanding these differences, we develop a chatbot-assisted self-regulated learning to create learning environments that support diverse student needs and ensure fair learning opportunities. Research studies various participants to determine whether the methods lead to successful outcomes for specific student groups. According to Manzanares et al. (2023), students achieve better academic success through these strategies, which enhance their motivation and engagement. It became our foundation for evaluating the effectiveness of the learning model and students' achievement levels in chatbot-assisted self-regulated learning of basic electrical engineering.

The world has witnessed the debut of generative AI systems in recent times. Technologies such as ChatGPT further underscore the need to explore how University students can achieve independent learning with essential technological support. This exploration shows that the situation worsens because most undergraduate students currently use AI-based chatbots for skill development, knowledge acquisition, and motivation. Ma et al.'s (2024) research shows that these intelligent agents' ability to adapt to individual learning paces and styles makes them invaluable tools for fostering proactive learning habits (Ng et al., 2024). Therefore, this study examines how AI-powered chatbots support self-regulated learning of basic electrical engineering concepts and assesses their impact on students with diverse backgrounds.

In line with that, our research questions are: RQ1: Is the basic electrical engineering chatbot-assisted self-regulated learning environment effective? RQ2: Do students' diversities have an impact on the learning model? RQ3: What is the students' perception of the usefulness and usability of the multi-role chatbot in facilitating their learning of basic electrical engineering concepts?

2. Literature Review

2.1. Artificial Intelligence and Learning

The deployment of AI-based chat tools in educational settings leads to significant changes in students' interactions with materials and processes, requiring thorough evaluation against current educational theories (Bognár et al., 2024). AI chatbots deliver learning support that aligns with Vygotsky's Zone of Proximal Development theory by providing appropriate assistance to students at their current learning level (Hauser et al., 2024). Scaffolding systems help students develop an understanding and acquire new skills by connecting their current abilities to their future potential under AI guidance (Darvishi et al., 2023). The framework of Self-Determination Theory enables the analysis of how AI chatbots impact intrinsic motivation, relatedness, competence, and autonomy, which drive self-regulated learning (Chiu, 2024).

Meeting these psychological requirements enables chatbots to create an educational environment in which students direct their learning journey and perceive AI assistance rather than authority (McGrath et al., 2024). Research

supports the idea that interactive support systems with custom-configured AI chatbots reduce both internal and external mental workload and elicit positive emotional responses (Becker et al., 2025). AI chatbots generate customized educational content that keeps students focused while developing their problem-solving and innovative thinking skills (Wang & Xue, 2024; Xu et al., 2024). The systems provide instant feedback that adapts to each student's learning rate to achieve better educational outcomes (Chen, 2024).

Some research on the adoption of AI chatbot educational support requires identifying variables that influence students' acceptance of the technology (Roca et al., 2024). Evidence from a systematic review indicates that AI chatbots have become more common in educational environments, as advances in AI and natural language processing have expanded their capabilities (Lo et al., 2024). Generative AI chatbots have become more advanced, so scientists lack enough empirical studies that use proven learning theories to measure their effects on student learning and teaching methods. AI chatbots in education demonstrate the positive benefits of generative AI for personalized learning and student motivation (Wang & Li, 2024).

However, few studies provide empirical evidence of their use in electrical engineering applications. Further research is still needed to assess these tools using established pedagogical frameworks, including experiential, reflective, active, and self-regulated learning, as current studies do not. Information deficiency by studying the impact of a Self-Determination Theory-based multi-role chatbot on students' self-regulated learning behaviors and their basic electrical engineering academic performance is also crucial. It is necessary to examine, through scientific methods, whether GenAI systems with theoretical foundations lead to better educational outcomes than the initial positive user experiences with these tools.

2.2. Effectiveness of AI integration in learning

Academic professionals are increasingly interested in ChatGPT and similar generative AI chatbots, but further research is needed to establish their effects on student educational outcomes. Research indicates that AI tutoring systems and domain-specific chatbots yield positive learning outcomes (Gao et al., 2024; Liu et al., 2025). Nevertheless, as scientists increasingly study ChatGPT and similar generative AI chatbots, it is essential to conduct further studies to determine their effects on self-regulated learning in basic electrical engineering. The research addresses a knowledge gap by investigating how a multi-role chatbot system helps students develop self-regulated learning skills in basic electrical engineering, thereby achieving better academic results. The method follows the principles of Generative Learning Theory because it enables learners to build knowledge through active construction, a process that AI chatbots support through customized, interactive prompts, as Wu et al. (2024) said. The educational generative AI architecture uses Retrieval-Augmented Generation to combine pre-trained language models with information retrieval systems, which produce answers through contextual understanding. The integration provides the chatbot with precise, up-to-date information, thereby improving instructional quality, according to Asghar et al. (2025).

2.3. Basic electrical engineering learning outcome through using AI-generated

Consequently, it is imperative to examine the impact of these resources on students' engagement, comprehension, and application of theoretical concepts within the electrical engineering curriculum. This study will also assess the efficacy of AI-powered conversational agents in enhancing deep learning and problem-solving skills, which are crucial for comprehending complex engineering principles (Caccavale, 2025). It entails assessing whether the chatbot's interactive features yield measurable improvements in students' analytical competencies and their ability to apply theoretical knowledge to practical circuit analysis and design challenges.

To gain a comprehensive understanding of electrical engineering, the chatbot must be able to interpret and generate responses from multimodal inputs, including circuit diagrams and graphical representations of electrical phenomena (Polverini & Gregorcic, 2024). Also, it is important to know how the chatbot's flexible learning paths can accommodate different learning styles and speeds, helping each student learn the basics of electrical engineering as quickly as possible. The students understand how integrating advanced machine learning and natural language processing algorithms into the chatbot can provide real-time, customized feedback, thereby

reconciling the disparity between the acquisition of theoretical knowledge and the development of practical skills (Thway et al., 2024).

The personalized feedback is important for helping students learn and encouraging them to correct themselves, ultimately helping them better understand complex electrical engineering concepts. Furthermore, the chatbot's ability to foster adaptive expertise, enabling students to apply acquired problem-solving techniques to unfamiliar and ambiguous engineering contexts, is a competency frequently absent in conventional curricula. In electrical engineering, students often encounter circuit configurations unfamiliar to them, making it hard to recognize patterns (Ridgway & Cox, 2024). This ability to adapt and apply what they know is essential. The model was described in Figure 1.

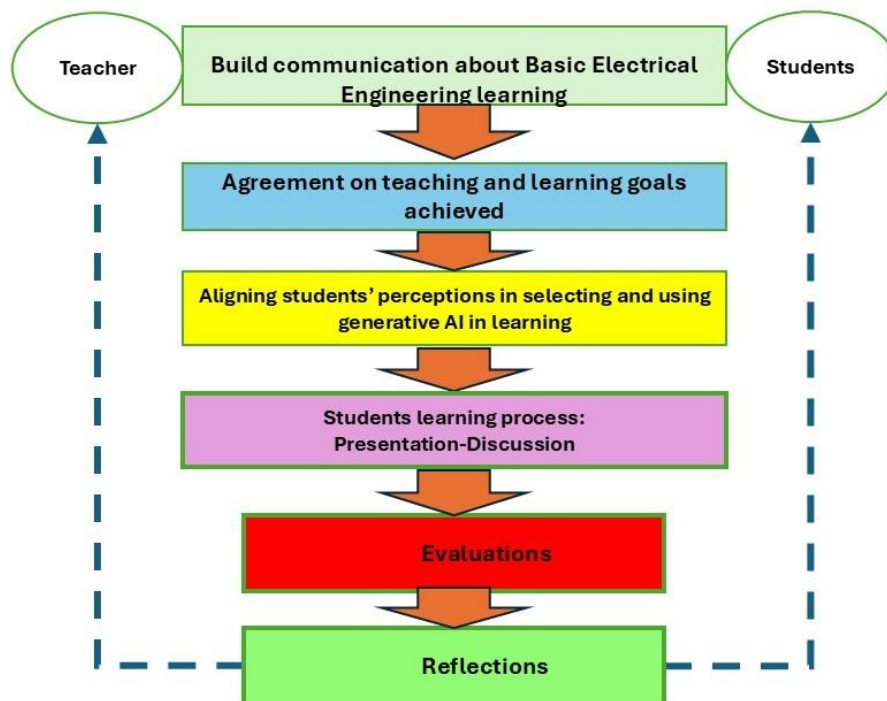


Figure 1: Chatbot-Assisted Self-Regulated Learning Model in Basic Electrical Engineering

3. Method

3.1. Learning model preparation

We created a learning model specifically for understanding basic electrical engineering by using large language models. We enable students to engage in deep, conversational interactions and use intelligent tutoring system features. The platform enables the development of complex AI-powered tools that deliver personalized learning materials and provide students with the feedback they need to encourage active learning and develop their critical thinking skills. We make the whole experience more user-friendly, so students can easily access learning materials and interact with teachers/staff on Microsoft Teams in a very comfortable, familiar way. The development process primarily involves building a detailed knowledge base from standard electrical engineering textbooks, research papers, and industry standards, ensuring the information provided is accurate and up to date.

3.2. Participants

We involved all students who took the Basic Electrical Engineering subject. There are six class groups (Group 1: 27 students; Group 2: 25 students; Group 3: 28 students; Group 4: 28 students; Group 5: 26 students; and Group 6: 25 students). Based on the census, the origin of school types is as follows: senior high school = 106 students;

vocational high school = 53 students; public school = 96 students; and private school = 63 students. There are 109 males and 50 females (Table 1).

Table 1: Group participants based on Subject Factors

		Value Label	N
School Status	1	Public School	96
	2	Private School	63
School Types	1	Senior High School	106
	2	Vocational High School	53
Gender	1	Male	108
	2	Female	51

3.3. Research design

We assess the participants' initial ability before implementing the learning through a pre-test and administer a post-test on the learning outcomes at the end of the instruction (Figure 2). We measure learning effectiveness using the Normalized Gain Rules (Hake, 1998). Investigating the impacts of participants' diverse backgrounds on learning effectiveness. Furthermore, we trace the participants' attitudes and perceptions toward learning.

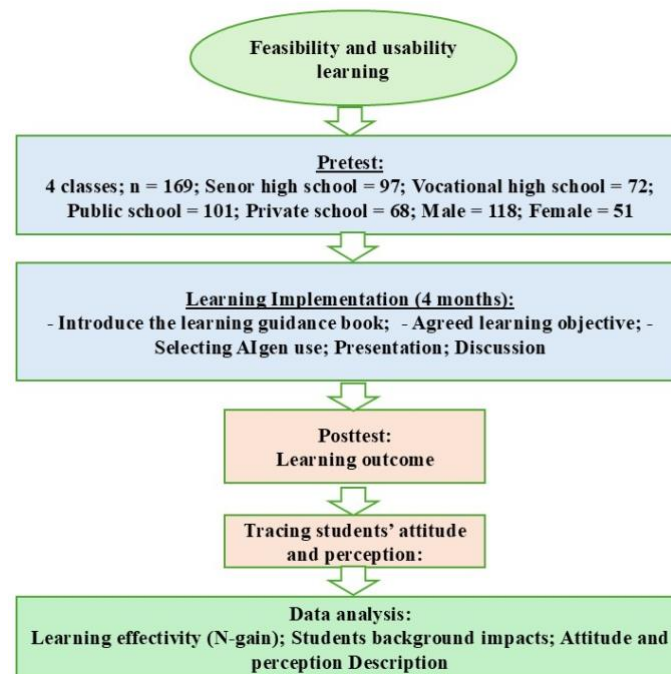


Figure 2: Research Design

3.4. Instruments and Data Collection

We first developed all the instruments and then tested their validity and reliability. There are three instruments: an assessment sheet to validate the model, involving the teacher, students, and the expert; tests to collect initial data on basic electrical engineering ability and learning outcomes; and Questionnaires to assess participants' attitudes and perceptions toward learning.

The assessment sheet consists of 25 valid items with a reliability coefficient of .71. The test consists of 40 valid items, including 30 multiple-choice and 10 essay items, with a reliability coefficient of .70. The assessment score is 0-100, with 2 points for correct multiple-choice answers and 4 points for correct essay answers. Questionnaires consist of 50 valid items (Likert Scale, 1 = Strongly disagree; 2 = Disagree ;3 = Enough agree; 4 = Agree; 5 = Strongly agree) with a reliability coefficient of .72. Minimum score = 50, maximum 250.

To validate the learning model, data were collected and analyzed in two rounds: Round 1 involved 20 students, one lecturer, and one expert; and Round 2, 55 students, three lecturers, and two experts. A pre-test was used to measure the students' initial abilities. Learning outcomes were collected after implementing the model for one semester, and students' perceptions of the learning model were also collected.

3.5. Data analysis

Quantitative data from pre- and post-intervention assessments were analyzed using Hake's rule to determine learning effectiveness,

$$g = \frac{\text{Post_test score} - \text{Pre_test score}}{\text{Maximum possible score} - \text{Pre_test score}} \quad (\text{Hake, 1998})$$

Description:

g = Normalized gain

Post-test score = the score on the post-test

Pre-test score = the score on the pre-test

Maximum possible score = the maximum possible score on the test = 100

Measuring criteria of learning effectiveness based on Hake, namely High gain = $g \geq .7$ (highly effective); Medium gain = $.3 \leq g < .7$ (moderately effective); and Low gain = $g < .3$ (less effective), and additionally, analyzing student learning outcome gain based on the students' school varieties by ANOVA at a .05 significance level. To describe students' attitudes and Perception trend levels regarding learning, descriptive statistics were used. The minimum ideal score is 50, and the maximum is 250. There are five categories of attitude and perception trend level determined: Very positive, $210 < \text{Mean} \leq 250$; Positive, $170 \leq \text{Mean} < 210$; Quite positive, $130 \leq \text{Mean} < 170$; Less favorable, $90 \leq \text{Mean} < 130$; Very less favorable, $50 \leq \text{Mean} \leq 90$ —all statistical analyses using IBM SPSS.

4. Results

4.1. Learning model feasibility

Based on the Round 1 validation, which involved 20 students, one lecturer, and one expert, several aspects need improvement, including the learning flow, the learning guide, the scope of the learning materials, and the measurement of learning objective achievement. After revising the model, it was implemented with 55 students and validated by three lecturers and two experts (Round 2). The students', three lecturers', and two experts' assessments concluded that the model was suitable (Table 2). It means the model is ready to use.

Table 2: Learning model feasibility test

Try out	Students	Lecturer	Expert	Conclusion
Round 1	20	1	1	Revisions needed: the learning flow, the learning guide, the scope of learning materials, and the measurement of learning objectives achievement.
Round 2	55	3	2	No revisions (feasible)

4.2. Learning model effectiveness

Mean scores of pre-tests, post-tests, and gain (Table 3) are the basis for determining the learning effectiveness. The normalized gain (g) calculation is $\frac{49.04}{66.54} = .737 > .7$. It describes a learning model that is highly effective in basic electrical engineering instruction.

Table 3: Pre-test score, post-test score, and learning outcome gain score descriptions

	N	Minimum	Maximum	Mean	Std. Deviation
Pre-test Scores	159	25	40	33.46	3.548
Post-test Scores	159	75	90	82.50	4.206
Outcome Gains	159	38	63	49.04	4.639

4.4. The gain in students' learning outcomes by Background.

Each group's student learning outcomes (gains) by school type, status, and gender (Table 4) showed that females outperformed males in all groups. The gain scores of participants' group learning outcomes serve as the basis for analyzing the role of students' existing diversity in learning.

Table 4: Participants' group gain score descriptions

Dependent Variable: Outcome Gains					
School Status	School Types	Gender	Mean	Std. Deviation	N
Public School	Senior High School	Male	47.95	4.274	38
		Female	50.57	5.372	21
	Vocational High School	Male	48.67	3.886	24
		Female	49.23	3.539	13
	Total	Male	48.23	4.111	62
		Female	50.06	4.741	34
Private School	Senior High School	Male	48.86	4.679	36
		Female	51.00	6.527	11
	Vocational High School	Male	48.60	4.006	10
		Female	49.83	5.742	6
	Total	Male	48.80	4.500	46
		Female	50.59	6.104	17
Total	Senior High School	Male	48.39	4.469	74
		Female	50.72	5.692	32
	Vocational High School	Male	48.65	3.860	34
		Female	49.42	4.194	19
	Total	Male	48.47	4.270	108
		Female	50.24	5.179	51

It was found that students' variety of school backgrounds and gender did not significantly affect the learning outcomes of chatbot-assisted self-regulated learning (Table 5). It was described significance of School Status (Sig. = .595 > .05), School type (Sig. = .562 > .05), and gender (Sig. = .065 > .05). Likewise, their fellow interactions did not impact (School status and School type, Sig. = .820 > .05; School status and Gender, Sig. .950 > .05; School type and gender, Sig. = .402 > .05; School status, School type, and gender, Sig. = .744 > .05).

Table 5: Tests of Between-Subjects Effects

Dependent Variable: Outcome Gains					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	147.556 ^a	7	21.079	.979	.449
Intercept	269601.001	1	269601.001	12517.536	.000
Sch_Status	6.105	1	6.105	.283	.595
Sch_Type	7.266	1	7.266	.337	.562
Gender	74.477	1	74.477	3.458	.065
Sch_Status * Sch_Type	1.125	1	1.125	.052	.820
Sch_Status * Gender	.059	1	.059	.003	.958
Sch_Type * Gender	15.218	1	15.218	.707	.402
Sch_Status * Sch_Type * Gender	2.306	1	2.306	.107	.744
Error	3252.218	151	21.538		
Total	385747.000	159			
Corrected Total	3399.774	158			

a. R Squared = .043 (Adjusted R Squared = -.001)

4.5. Students' Attitudes and Perceptions on the learning model

The trend in students' attitudes and perceptions of learning, based on school status, school type, and gender, is positive. Their means are in $170 \leq \text{Mean} < 210$ (Table 6). It describes that the students tend to accept the learning model positively. Based on interviews with several students, they generally said the learning model facilitates their learning.

Table 6: Students' Attitude and Perception Description

		N	Mean	Std. Deviation
Attitude and Perception	All	159	181.85	18.103
School status	Public School	96	180.70	18.715
	Private School	63	183.60	17.128
School type	Senior High School	106	181.85	17.852
	Vocational High School	53	181.85	18.770
Gender	Male	108	181.68	18.136
	Female	51	182.22	18.208

5. Discussion

5.1. Chatbot-assisted self-regulated learning environment effectiveness

The developed learning model proved effective for use in basic electrical engineering instruction. This effectiveness was highlighted by its ease of use and practicality. However, it is important to emphasize that adaptive learning tools that accurately measure student understanding and provide the necessary support to correct misconceptions remain needed (Afzaal et al., 2021). Specifically, examining students' behavioral patterns when interacting with the chatbot could enable the computer to adjust teaching methods more quickly, as Caccavale (2025) suggested.

The insights gained from these exchanges inform the creation of more individualized learning paths, enabling the chatbot to guide students through remedial modules or offer additional challenges based on their demonstrated skills and areas of weakness. Furthermore, these interaction patterns can reveal the effectiveness of various chatbot

responses, thus encouraging incremental adjustments to the chatbot's pedagogical design and its overall effectiveness in supporting self-regulated learning (Lin et al., 2024). Such an iterative improvement process, guided by continuous review of user interactions and learning outcomes, is crucial for intelligent tutoring systems to effectively promote self-regulation (Ng et al., 2024; Mejeh et al., 2024). Research has revealed that some students exhibit overconfidence in their understanding, particularly when rejecting chatbot guidance on complex problems, which can lead to incorrect corrections. This fact highlights the need to develop sophisticated AI mechanisms that can detect such metacognitive illusions and prevent them through adaptive feedback and corrective prompts (Kumar et al., 2024).

5.2. Students' diversity's impact on learning

The findings from this study suggest that various demographic factors, such as school type, socioeconomic status, and gender, do not significantly influence learning outcomes within chatbot-assisted self-regulated learning environments. This outcome diverges from conventional educational research that frequently identifies these variables as significant predictors of academic achievement in traditional settings. Instead, the consistency observed across diverse student backgrounds highlights the potential of chatbot-assisted self-regulated learning to offer an equitable and universally effective pedagogical approach (Ng et al., 2024). This observed uniformity suggests that the scaffolding provided by educational chatbots in facilitating self-regulated learning processes may effectively mitigate pre-existing disparities arising from varied educational or socioeconomic contexts (Guan et al., 2024).

The mitigation underscores the capability of AI-driven tools to democratize access to effective learning strategies, enabling students from all backgrounds to develop critical metacognitive skills and improve academic performance (Chang et al., 2023). This aligns with broader research indicating that chatbot technology can exert a medium-to-high effect on overall learning outcomes, independent of such moderator variables (Deng & Yu, 2023). Specifically, metacognitive support delivered via chatbots has been shown to enhance self-regulated learning abilities, particularly in task strategy and self-evaluation, thereby optimizing the learning experience for all students (Xu et al., 2025). This further suggests that the inherent adaptability and personalized feedback mechanisms embedded in chatbot technologies effectively cater to individual learning needs, thereby neutralizing the differential impact typically associated with diverse student demographics (Manzanares et al., 2023).

5.3. Students' perception of the learning's usefulness and usability

Students generally viewed the chatbot as an effective tool for personalized learning and quick feedback, which in turn would lead to a more efficient grasp of the course material, in line with Chen (2024). Nevertheless, this favorable reception did not extend to all aspects of the chatbot's functionality, as some students doubted its ability to handle complex or subtle inquiries.

The need for the chatbot to have adaptive response depth, which could adjust the complexity of the answer to the student's level of expertise, was repeatedly raised, suggesting that students wanted more intricate interactions. Moreover, examination of the data concerning the students' opinions about the chatbot revealed that their principal appreciation of the chatbot was to be used for clarification of doubts and understanding of concepts; however, they considered the chatbot to be less helpful in the areas of thorough content review and exam simulations (Vera, 2024). Although the chatbot successfully served as an on-demand tutor for certain concepts, its current configuration may not fully support broader learning goals or accommodate diverse learning strategies.

This highlights the need for future versions to include stronger systems for more thorough review, possibly through structured learning paths or connections to external knowledge, to increase their overall educational value (Lin et al., 2024). The chatbot was also seen as having potential to promote collaboration and collaborative learning, primarily through collaborative problem-solving and peer explanation. Moreover, features that allow students to record their thought processes or share interpretations with a virtual peer would foster higher-order thinking and understanding beyond mere memorization.

5.4. School background and demographic factors mediate effectiveness and engagement with a chatbot-assisted self-regulated learning.

Students' backgrounds and demographic factors, prior academic achievement, and learning styles may substantially influence their interactions with and perceptions of the chatbot, suggesting that a one-size-fits-all approach to integrating AI into education may not be practical (Stöhr et al., 2024). It also reiterates the need for personalized chatbot design that adapts to diverse learning profiles, thereby enhancing engagement and learning outcomes across a diverse student population (Sandu et al., 2024).

For example, students who prefer visual learning would benefit from chatbots with interactive diagrams and simulations in their interfaces. In contrast, students who prefer auditory learning benefit more from chatbots with audio explanation features. Students with a preference for kinesthetic learning may benefit from chatbot features that incorporate virtual experiments or active problem-solving scenarios (Sadegh-Zadeh et al., 2023).

The effectiveness of these individualized approaches is further corroborated by findings that students' prior academic achievement was a significant predictor of how quickly they used the chatbot and of their performance improvement. For instance, students with high prior attainment frequently used chatbot assistance to ask advanced questions and broaden their knowledge. In contrast, students with lower prior attainment tended to use it only to build their foundational knowledge (Wambsganß et al., 2024). In contrast, students with lower prior academic achievement often require more structured, guided interactions with chatbots, relying on them as a primary resource to clarify key concepts and help address their near-complete lack of knowledge.

The findings point to the necessity of adaptive scaffolding tailored to individual academic experiences and learning preferences as a prerequisite for increasing chatbot efficacy and ensuring equitable access to their learning potential (Hauser et al., 2024). Differentiating support at the individual level through more granular learner profiling can improve the chatbot's utility for a broader range of students. Doing so would involve implementing advanced diagnostic techniques embedded in the chatbot to assess students independently and dynamically update their instructional approach (Stöhr et al., 2024). Having insight into user preferences and choices, including the number of questions or the time required to achieve the outcome, can enable the chatbot to engineer prompts that best suit individual student preferences and improve the efficiency of their learning experience.

6. Conclusion

The efficacy of chatbots depends on their personalized and scaffolding capabilities, which enhance reading comprehension and problem-solving, especially within a Retrieval-Augmented Generation architecture. The adaptability of this approach allows the chatbot to provide scaffolded responses and thorough steps, which are essential for constructing knowledge and developing understanding.

7. Limitation

Despite the promising findings, this study encountered several limitations that warrant consideration for future research, including the relatively short study horizon and the focus on a single academic discipline, which may limit the generalizability of the results.

Author Contributions: All authors contributed to this research.

Funding: Universitas Negeri Medan, Grant Number: 01994/UN33/KPT/2023.

Conflict of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics Approval: Not applicable.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., Li, X., & Weegar, R. (2021). Explainable AI for Data-Driven Feedback and Intelligent Action Recommendations to Support Students' Self-Regulation. *Frontiers in Artificial Intelligence*, 4. <https://doi.org/10.3389/frai.2021.723447>
- Asghar, M. Z., Hashmi, Z. F., Asghar, M. Z., & Abid, M. N. (2025). The use of generative AI tools enhances academic achievement in sustainable education by fostering shared metacognition and cognitive offloading among preservice teachers. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-01676-x>
- Bognár, L., Ágoston, G., Bacsa-Bán, A., Fauszt, T., Gubán, G., Joós, A., Juhász, L., Kocsó, E., Kovács, E., Maczó, E. C., Kollár, A. I. M., & Strauber, G. (2024). Re-Evaluating Components of Classical Educational Theories in AI-Enhanced Learning: An Empirical Study on Student Engagement. *Education Sciences*, 14(9), 974. <https://doi.org/10.3390/educsci14090974>
- Caccavale, F. (2025). From Programming to Chatbots: Digitalization Strategies in Chemical Engineering Education. *Research Portal Denmark*, 189. <https://local.forskningsportal.dk/>
- Chang, D., Lin, M. P.-C., Hajian, S., & Wang, Q. Q. (2023). Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization. *Sustainability*, 15(17), 12921. <https://doi.org/10.3390/su151712921>
- Chen, W. (2024). Intelligent Tutor: Leveraging ChatGPT and Microsoft Co-Pilot Studio to Deliver a Generative AI Student Support and Feedback System within Teams. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2405.13024>
- Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: a case of ChatGPT. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-024-10366-w>
- Deng, X., & Yu, Z. (2023). A Meta-Analysis and Systematic Review of the Effect of Chatbot Technology Use in Sustainable Education. *Sustainability*, 15(4), 2940. <https://doi.org/10.3390/su15042940>
- Gao, Z., Cheah, J., Lim, X., & Luo, X. (2024). Enhancing academic performance of business students using generative AI: An interactive-constructive-active-passive (ICAP) self-determination perspective. *The International Journal of Management Education*, 22(2), 100958. <https://doi.org/10.1016/j.ijme.2024.100958>
- Guan, R., Raković, M., Chen, G., & Gašević, D. (2024). How do educational chatbots support self-regulated learning? A systematic review of the literature [Review of *How educational chatbots support self-regulated learning? A systematic review of the literature*. *Education and Information Technologies*. Springer Science+Business Media. <https://doi.org/10.1007/s10639-024-12881-y>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Hauser, A. S., Hansen, U., & Skov, S. (2024). AI chatbots as Co-Pilots in Learning During Practical Teaching Activities. *Research Portal Denmark*, 2. <https://local.forskningsportal.dk/>
- Lan, M., & Zhou, X. (2025). A qualitative systematic review of AI-enabled self-regulated learning in higher education. *Npj Science of Learning*, 10(1). <https://doi.org/10.1038/s41539-025-00319-0>
- Lin, M. P.-C., Chang, D., & Winne, P. H. (2024). A proposed methodology for investigating student-chatbot interaction patterns in giving peer feedback. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-024-10408-3>
- Lin, X., Luterbach, K., Gregory, K., & Sconyers, S. E. (2024). A Case Study Investigating the Utilization of ChatGPT in Online Discussions. *Online Learning*, 28(2). <https://doi.org/10.24059/olj.v28i2.4407>
- Lo, C. K., Hew, K. F., & Jong, M. S. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda [Review of *The influence of ChatGPT on student engagement: A systematic review and future research agenda*]. *Computers & Education*, 219, 105100. Elsevier BV. <https://doi.org/10.1016/j.compedu.2024.105100>
- Ma, W., Wen-jing, M., Hu, Y., & Bi, X. (2024). The who, why, and how of AI-based chatbots for learning and teaching in higher education: A systematic review [Review of *The who, why, and how of AI-based chatbots for learning and teaching in higher education: A systematic review*]. *Education and Information Technologies*. Springer Science+Business Media. <https://doi.org/10.1007/s10639-024-13128-6>
- Makransky, G., Plechatá, A., Herlau, T., & Blurton, S. P. (2025). Beyond the “Wow” Factor: Using Generative AI for Increasing Generative Sense-Making. *Educational Psychology Review*, 37(3). <https://doi.org/10.1007/s10648-025-10039-x>

- Manzanares, M. C. S., Sánchez, R. M., Antón, L. J. M., González-Díez, I., & Almeida, L. S. (2023). Perceived satisfaction of university students with using chatbots as a tool for self-regulated learning. *Heliyon*, 9(1). <https://doi.org/10.1016/j.heliyon.2023.e12843>
- McGrath, C., Farazouli, A., & Pargman, T. C. (2024). Generative AI chatbots in higher education: a review of an emerging research area [Review of *Generative AI chatbots in higher education: a review of an emerging research area*]. *Higher Education*. Springer Science+Business Media. <https://doi.org/10.1007/s10734-024-01288-w>
- Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, 55(4), 1328. <https://doi.org/10.1111/bjet.13454>
- Polverini, G., & Gregorcic, B. (2024). Evaluating vision-capable chatbots in interpreting kinematics graphs: a comparative study of free and subscription-based models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2406.14685>
- Ridgway, L. M., & Cox, T. (2024). Investigating Student Approaches to Rearranging Circuit Diagrams. *IEEE Transactions on Education*, 67(5), 681. <https://doi.org/10.1109/te.2024.3410375>
- Roca, M. D. L., Morales, M., García-Cabot, A., García-López, E., & Amado-Salvatierra, H. R. (2024). The impact of a chatbot working as an assistant in a course for supporting student learning and engagement. *Computer Applications in Engineering Education*, 32(5). <https://doi.org/10.1002/cae.22750>
- Sadeh-Zadeh, S., Movahhedi, T., Hajiyavand, A. M., & Dearn, K. D. (2023). Exploring undergraduates' perceptions of and engagement in an AI-enhanced online course. *Frontiers in education*, 8. <https://doi.org/10.3389/educ.2023.1252543>
- Sandu, R., Gide, E., & Elkhodr, M. (2024). The role and impact of ChatGPT in educational practices: insights from an Australian higher education case study. *Discover Education*, 3(1). <https://doi.org/10.1007/s44217-024-00126-6>
- Stöhr, C., Ou, A. W., & Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels, and fields of study. *Computers and Education Artificial Intelligence*, 7, 100259. <https://doi.org/10.1016/j.caeai.2024.100259>
- Thway, M., Recatala-Gomez, J., Lim, F. S., Hippalgaonkar, K., & Ng, L. W. T. (2024). Harnessing GenAI for Higher Education: A Study of a Retrieval Augmented Generation Chatbot's Impact on Human Learning. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2406.07796>
- Vera, F. S. (2024). Subject-Specialized Chatbot in Higher Education as a Tutor for Autonomous Exam Preparation: Analysis of the Impact on Academic Performance and Students' Perception of Its Usefulness. *Education Sciences*, 15(1), 26. <https://doi.org/10.3390/educsci15010026>
- Wambsganß, T., Benke, I., Maedche, A., Koedinger, K. R., & Käser, T. (2024). Evaluating the Impact of Learner Control and Interactivity in Conversational Tutoring Systems for Persuasive Writing. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-024-00409-x>
- Wang, Y., & Xue, L. (2024). Using AI-driven chatbots to foster Chinese EFL students' academic engagement: An intervention study. *Computers in Human Behavior*, 159, 108353. <https://doi.org/10.1016/j.chb.2024.108353>
- Wu, D., Zhang, S., Ma, Z., Yue, X., & Dong, R. K. (2024). Unlocking Potential: Key Factors Shaping Undergraduate Self-Directed Learning in AI-Enhanced Educational Environments. *Systems*, 12(9), 332. <https://doi.org/10.3390/systems12090332>
- Xu, Y., Zhu, J., Wang, M., Qian, F., Yang, Y., & Zhang, J. (2024). The Impact of a Digital Game-Based AI Chatbot on Students' Academic Performance, Higher-Order Thinking, and Behavioral Patterns in an Information Technology Curriculum. *Applied Sciences*, 14(15), 6418. <https://doi.org/10.3390/app14156418>
- Yin, J., Xu, H., Pan, Y., & Hu, Y. (2025). Effects of different AI-driven Chatbot feedback on learning outcomes and brain activity. *Npj Science of Learning*, 10(1). <https://doi.org/10.1038/s41539-025-00311-8>

Assessing Instructional Leadership Practices on Enhancing Pre-Service Teachers' Teaching Competency at Battambang Teacher Education College, Cambodia

Sopheaktra Chea¹, Chan Hum², Sam Rany³, Merryl Kravitz⁴

¹ Graduate School, National University of Battambang, Cambodia. Email: cheasopheaktra5678@gmail.com

² Graduate School, National University of Battambang, Cambodia. Email: hum.chan@nubb.edu.kh

³ Graduate School, National University of Battambang, Cambodia. Email: sam.rany@nubb.edu.kh

⁴ Professor of Secondary Education at New Mexico Highlands University, linkedin.com/in/merryl-kravitz-581913123

Correspondence: Merryl Kravitz, New Mexico Highlands University

Abstract

This study analyzes the relationship between instructional leadership practices and the development of teaching competency among pre-service teachers at the Battambang Teacher Education College (BTEC) in Cambodia. Although instructional leadership is recognized as a primary factor in teacher effectiveness worldwide, empirical evidence of its practice in the Cambodian teacher education context remains scarce. The study used a mixed-methods approach and collected data from 365 participants, including 303 pre-service teachers, 21 instructors, 31 student-teachers, and 10 management committee members. The findings revealed a substantial link between leader training and teachers' skills. The research also revealed that guidance from experienced instructors truly helped new teachers improve their instructional skills. Providing feedback and training for new teachers is also crucial. However, the study had an issue. The schools new teachers go to for trial teaching don't always have what they need. Those responsible need to ensure they support the instructors. The report states that strong leadership is required to ensure that pre-service teachers are well equipped. So, this study's results should be used to improve teacher training programs by the people running them." Teacher education institutes should use this information to learn from how they train instructors, helping teachers do their jobs better.

Keywords: Teaching Competency, Instructional Leadership, Pre-Service Teachers, Battambang Teacher Education College

1. Introduction

In Cambodia, pre-service teachers need to receive the same quality of education as students in other countries. Battambang Teacher Education College supports this by helping pre-service teachers train and build the skills they need to teach effectively. Battambang Teacher Education College helps ensure teachers are prepared. Teaching skills are essential for teachers. In this way, teachers can help Cambodian students learn. Cambodian students need

quality teachers. Battambang Teacher Education College is helping to make this happen. This study examines the impact of instructional leadership on pre-service teachers' professional development and assesses how these leadership practices facilitate growth and pedagogical competence in a BTEC environment. The quality of an education system depends largely on the quality of its teachers. This puts a responsibility on teacher training institutions. They need to ensure pre-service teachers gain the knowledge and skills to do their job well.

To improve teaching and learning, educational administrators use various methods. They create plans for their schools. They design curricula that cover everything students need to learn. They provide mentorship and guidance to help pre-service teachers improve. They work to create a positive learning environment. Teachers trained at BTEC will go on to teach students for years. This makes BTEC very important for the country's education system. BTEC supports the education of many generations of students. The government and the college are working together to elevate educational standards; therefore, BTEC's effectiveness is vital to improving the Cambodian education system. Instructional leadership is recognized as a key determinant of pre-service teacher effectiveness.

This research aims to do the following: analyze the teaching leadership practices used by current BTEC teachers and administrators. Assess how well pre-service teachers learn to teach during their training. We want to see whether the way teachers are guided helps them become effective teachers. Good leaders can affect how well teachers teach and how many students learn. This study aims to identify what makes pre-service teachers effective and how to better engage students. We are exploring how leaders can help pre-service teachers improve. It helps us better understand teacher education in Cambodia than previous studies have. The study provides us with real-life examples of how instructional leadership works in Cambodian schools. This is useful for teacher education institutions because it gives them a blueprint for improving how teachers teach and how students learn. They can use the study's findings to inform decisions about improving pre-service teacher training in Cambodia. The study's findings can help those working on the 2024-2028 Education Strategic Plan for Cambodia make decisions about education in Cambodia. It examines the challenges and needs of Cambodian schools and provides practical solutions. The study's results can help teacher education institutions and policymakers work together to improve education in Cambodia. The primary focus is measuring pedagogical efficacy, specifically pre-service teachers' ability to plan lessons and implement instruction effectively in the classroom. Teaching skills are essential for teachers. In this way, teachers can help Cambodian students learn. Cambodian students need quality teachers. Battambang Teacher Education College is helping to make this happen. This study examines the impact of instructional leadership on pre-service teachers' professional development and assesses how these leadership practices facilitate growth and pedagogical competence in a BTEC environment. The quality of an education system depends largely on the quality of its teachers. This puts a responsibility on teacher training institutions. They need to ensure pre-service teachers gain the knowledge and skills to do their job well. To improve teaching and learning, educational administrators use various methods. They create plans for their schools. They design curricula that cover everything students need to learn. They provide mentorship and guidance to help pre-service teachers improve. They work to create a positive learning environment. Teachers trained at BTEC will go on to teach students for years. This makes BTEC very important for the country's education system. BTEC supports the education of many generations of students. The government and the college are working together to elevate educational standards; therefore, BTEC's effectiveness is vital to improving the Cambodian education system. Instructional leadership is recognized as a key determinant of pre-service teacher effectiveness.

This research aims to do the following: analyze the teaching leadership practices used by current BTEC teachers and administrators. Assess how well pre-service teachers learn to teach during their training. Determine whether the way teachers are guided helps them become effective teachers. Good leaders can make a difference in how well teachers teach and how many students learn. This study examines what makes pre-service teachers effective and how to engage students more. It explores how leaders can help pre-service teachers improve. It helps us better understand teacher education in Cambodia than previous studies have. The study provides us with real-life examples of how instructional leadership works in Cambodian schools. This is useful for teacher education institutions because it gives them a blueprint for improving how teachers teach and how students learn. The findings can inform decisions about how to improve pre-service teacher preparation in Cambodia. This study's conclusions can help those involved in the 2024-2028 Education Strategic Plan for Cambodia make decisions about education in Cambodia. It addresses schools' problems and needs in Cambodia and proposes practical solutions. Teacher education institutes and policymakers can use the findings of this study to work together to

improve education in Cambodia. The major concern is the measurement of pedagogical effectiveness, that is, the pre-service teachers' competence in lesson design and successful implementation of instruction in the classroom.

2. Literature Review

The literature review combines theoretical and empirical foundations relevant to instructional leadership and pre-service teacher competency. This study draws on educational theories that describe how leadership influences teacher performance. Hallinger (2011) reported that school administrators help develop a learning environment by directing instructional programs, aligning curricula, and monitoring student and teacher development. The study also draws on Bandura's (1977) theory of "learning," in which rookie teachers learn to teach by watching and imitating experienced mentors. Leadership is critical to ensuring instructors educate well and maintain standards. The study shows that school administrators build an environment that helps teachers do their best. It is defined as a forward-looking process in which leaders align institutional efforts to support student success through curricular management and collaborative professional development. Hallinger's review of four decades of research on effective leaders highlights three core functions of effective school leaders: setting the mission, administering instructional programs, and creating positive school climate. Linda Darling-Hammond (2006) found that ongoing support and formative feedback are critical to reducing new-teacher turnover and accelerating professional growth. Building on Stronge's Teacher Competency Framework, this study assesses pre-service teachers' competencies in lesson planning, effective instruction, and classroom management. Instructional leadership provides the framework that the pre-service teachers need to acquire these fundamental competencies. The existing literature largely focuses on educational environments in industrialized nations, leaving a considerable gap regarding the implementation of instructional leadership in the Cambodian teacher education system. This study addresses these shortcomings by examining how instructional leadership can empower pre-service teachers in the specific context of resource-limited BTEC. The factors of the study are analyzed by using the 'Teacher Competency Framework' (Khan, 2018) and 'Instructional Leadership Theory' (Hallinger, 2011). These strategies include classroom observation, formative feedback, professional development seminars, peer mentoring, and curricular support.

3. Methodology

The study focused on teaching skills and leadership practices in the Battambang Teacher Education College (BTEC). The study used a mixed-methods approach to provide a complete picture. The study used a sequential explanatory mixed-methods approach, collecting quantitative data followed by qualitative in-depth interviews. The study sample comprised 365 BTEC examinees. The sample comprised 303 prospective teachers. A stratified random sample was used to represent different academic specializations (e.g., Elementary Education, Science, and Khmer Literature) and years of study. The study also used purposive sampling, selecting 31 participants. The study included ten management team members and twenty-one teacher educators selected for their direct involvement in leadership and mentorship. These were the ones selected for this. Thirty-one pre-service teachers also participated in focus discussions to share their opinions on what they learned. This survey used a five-point scale, allowing respondents to indicate whether they strongly disagreed or strongly agreed. The survey was based on the Teaching Competency Scale, adapted from Stronge's Teacher Competency Framework (2018) and Darling-Hammond (2006). It also used the Instructional Leadership Scale, based on the work of Hallinger (2011). The researchers processed the statistical data using SPSS. The mean and standard deviation summarized participants' responses. The focus group transcripts were evaluated using the six-step framework developed by Braun and Clarke. The survey used the Teaching Competency Scale and the Instructional Leadership Scale. The analysis involved data familiarization, categorization, and systematic identification of recurring themes related to leadership and instructional efficacy.

4. Results

The study is based on an analysis of data received from 303 pre-service teacher respondents from the Battambang Teacher Education College (BTEC). The quantitative data were analyzed using statistical tools to identify the association between instructional leadership techniques and teacher competency. The qualitative data added

contextual explanations to the findings. The average scores of instructional leadership techniques were used to assess the effectiveness of current leadership initiatives. The statistics showed that pre-service teachers fell along a continuum of teaching skills, with some areas more developed than others. Correlation Coefficient ($r = 0.538$): This shows a strong positive relationship between instructional leadership and the teaching competency of pre-service teachers. The higher the quality of instructional leadership, the higher teaching competency tends to be. The statistical study shows a very substantial association ($p < 0.001$), meaning the correlation between instructional leadership and teaching competency is meaningful and not accidental. What is interesting is the very low p-value. So, we can be sure that the connection between these items is real and very unlikely to be a coincidence. Teachers partnered with leaders who value teaching, teamwork, and offering positive feedback are often particularly successful at preparing lessons, conducting courses, and managing the classroom. Teachers who work with these kinds of leaders are better at preparing lessons, giving instruction, and managing the classroom. This is where the Significance Level becomes important, as it shows there is a statistically significant relationship between instructors and leaders who focus on instructional modeling, team cooperation, and constructive criticism.

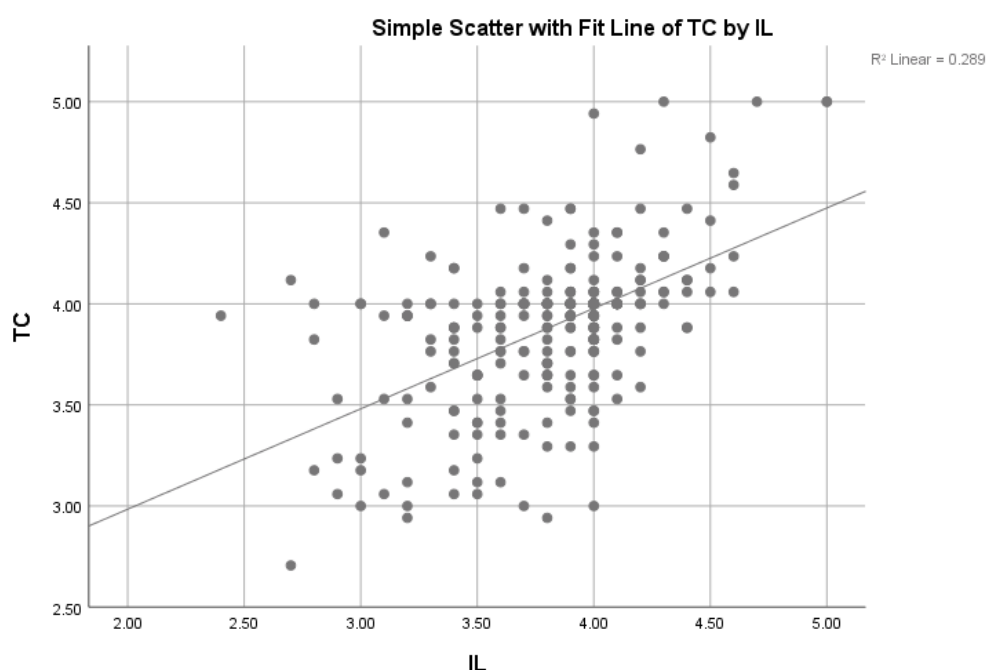


Figure 1: Scatter Plot of Instructional Leadership and Teacher Competency

IL is plotted on the x-axis and TC on the y-axis. IL increases as TC increases. This is shown by the upward-sloping line. The $R^2 = 0.289$. This means changes in IL can explain 28.9% of the changes in TC. The remaining 71.1% comes from other factors or random variation. The overall trend is upwards, but the dots are scattered along the line. This shows some variation. Also, near the top of TC, where TC equals 5.00, there is a limiting effect.

4.1. Enhancing Instructional Leadership for Student Teachers

Findings

Good leadership connects and directs people. Good leadership isn't about giving orders from an administrative position; it is about being involved and providing a collaborative direction. Good leadership fosters an environment where people work to help and learn from one another. It helps individuals feel valued and supported. It creates trust and great partnerships. The leader has to show others the way. This includes actions for the school and for the people who work there. BTEC is struggling with infrastructural issues, such as a lack of technology and budget limits. It also lacks formal leadership training. The Social Learning Theory (Bandura, 1977) illustrates this in relation to instructional frameworks. "We can't do it alone. If we want to succeed, we need to work as a team. We should make sure the individuals in control get the training they need, everyone shares what they know, we employ

technology and teaching methods, and that the leaders have the time and power to make sure everything goes properly.

Students and leaders must develop their skills and improve their practice. They need to learn from each other.” Student instructors and leaders should ask for help when they need it. The leaders should be able to pay attention to their work. They have to ensure student teachers and leaders are doing everything right. When student instructors and leaders do this, they will do well in their roles. The tactics range from formal feedback and coaching to experience. We need these to make things happen. It should balance leadership development, resource allocation, and ongoing training. The best strategy to improve teacher effectiveness is one-on-one coaching and feedback. But it only works with help. Good leadership requires time, training, and resources to develop and support teachers and to build collaboration. Leadership is power. Good leaders lead, coach, and show teachers how to connect theory and practice. Leadership style is adaptable. Includes everyone. It gives structure and support, but also freedom for educators to flourish. They define curriculum, mentor teachers, and monitor instructional methods to meet school goals. But implementation is typically tricky. Two issues: Inadequate time, budget, resources, and infrastructure. Limitations: Training, ability, and motivation for students/faculty; faculty ability and motivation for students/faculty. We must help teachers develop and grow to improve leadership in schools. Methods for supporting teacher growth include mentoring, classroom observation with feedback, hands-on learning, and project-based collaboration. These work best when teachers also reflect on their own practice.

The Power of Leaders to Transform: “Leaders develop teachers by coaching them, by modeling for them, and by involving them in professional learning.” New teachers have two challenges. One is organizational and logistical: not having time, having too many pupils, or lacking the requisite resources or infrastructure. Another is personal: feeling insecure, struggling with lesson preparations, and having trouble keeping pupils on task. The Role of Leadership: Good leadership can help build confidence and competence in new instructors. Feedback and a supportive environment can help new instructors move from simply delivering a lecture to active, reflective teaching. This readies teachers for self-directed teaching. Instructional leadership at BTEC links theory and practice. Leaders who do more than manage, but who intentionally coach, give feedback, and foster a collaborative environment where new teachers can explore, reflect, and develop, get the best results. BTEC’s current instructional leadership techniques include a distributed team leadership style that fosters close cooperation among administrators, instructors, and support staff to advance institutional goals. As administrators, educators, and staff, we work together to set teaching goals and a shared vision. We work with schools and the community to ensure our training meets real needs in our area. This helps keep our training current.

Pedagogy: We engage learners through talks, group work, and practical classes. These strategies assist pupils to learn better. We use classroom observations, grading rubrics, and feedback to monitor teaching quality and hold teachers accountable for their work. We facilitate and make administrative and teaching responsibilities more accessible through methods such as virtual classrooms and online planning. Levels of Teaching Competency of Pre-Service Teachers: BTEC Competence represents a shift from knowledge transfer to learning facilitation. It is mapped over three levels. Instructional leadership ensures teachers can transfer their learning to their work. Leaders lead by example for instructors. When leaders model instructional techniques and technology, instructors are inspired to become more creative and to employ technology in their instruction. Instructional leadership provides teachers with a setting. It develops a team of teachers that collaborate and learn from each other. When teachers collaborate to develop curricula and teach, they share responsibility for students' success. Instructional leadership helps teachers work as a team.

Strategies for Improvement

Need to do something to make BTEC a better place for teachers to learn and grow. “Going to teach teachers to be strong leaders, to use technology in the classroom, to teach in innovative and exciting ways. We'll pair teachers with experienced teachers so they may learn from them and get solid suggestions. Innovation and Research will inspire instructors to try things, attend conferences, and share ideas with others. We need to do more to make BTEC a strong place for instructors to learn. Here's how we suggest these techniques will work. BTEC is doing a good job. They have a broad strategy to support teachers from the start of their careers through to leadership. This

transforms how the institution functions; it is no longer about delivering information to people, but about working together and helping each other flourish. People who go to BTEC will be good at what they do; they will also be able to reflect on their work, adapt as new challenges arise, and solve the complex problems education faces today. BTEC uses a technique with different aspects; thus, those in charge of teaching are not only administrators but also teachers, mentors, and people who are still learning to be teachers. This generates a cycle of greatness at BTEC that will go on.”

5. Discussion

The study shows that leadership style in teaching affects the teaching abilities of pre-service teachers at Battambang Teacher Education College (BTEC). The results suggest a link between a leader’s teaching style and teachers’ success in instructing their students. This connection is very strong. This means leaders are responsible for about half of how well teachers do their jobs. The Role of Instructional Leadership. The study confirms that instructional leadership is not merely an administrative function but a fundamental catalyst for teacher development. This finding aligns with Hallinger’s (2011) Instructional Leadership Theory, as it suggests that when BTEC administrators focus on managing instructional programs and providing positive feedback, pre-service teachers show higher competence in lesson planning and implementation. This confirms the premise that leadership provides a structural framework for instructional growth. Mentorship and social cognitive development: The qualitative results show the “observational professionalism” pre-service teachers experienced. In line with Bandura’s (1977) Social Cognitive Theory and Darling-Hammond’s (2006) work, this confirms that the relationship between leadership and competency is mutual. While leadership guides the pre-service teacher, active engagement in feedback cycles allows leaders to refine their supervisory practices. The problem is that what student teachers learn at BTEC differs from what they can actually do in schools. This is called the resource gap. When student teachers go to schools, they often find that the schools do not have the technology and resources that they learned about at BTEC. Therefore, school authorities should support student teachers in understanding how to teach in these scenarios. The resource gap is an issue that needs to be addressed. At BTEC, they teach student teachers how to teach, but then they go to schools that do not have the same resources. This makes it hard for the student teachers to use what they learned at BTEC. Those in charge need to develop plans to help student teachers address the resource gap. Implications for Policy and Practice: This study’s results have significant implications for Cambodia’s 2024-2028 Education Strategic Plan. To improve teacher quality, BTEC is advised to codify mentorship as a function and include Stronge’s Teacher Competency Framework in periodic supervisory assessments. Shifting leadership focus from administrative oversight to targeted instructional support better positions institutions to equip pre-service teachers to handle the challenges of the modern classroom.

6. Conclusion

This study provides an in-depth review of the influence of instructional leadership on the teaching capability of pre-service teachers at Battambang Teacher Education College (BTEC). The findings show that instructional leadership significantly affects teacher growth in Cambodia. There is a relationship between what instructional leaders do and how good teachers are at their work. This suggests that leadership quality contributes to how well teachers teach. The research determined that for teachers still learning, comments and observations of experienced teachers’ work are quite significant. This assists teachers learning the craft to see where they are headed. Teachers who are still learning need to see how experienced teachers work. The research also says that schools that train teachers need to prepare them, because many schools lack resources. Teachers who are still learning will work in schools that do not have many resources. The research says that schools that train teachers should prepare teachers for this. This way, teachers can apply what they learn at BTEC in real situations.

To summarize, the study suggests that instructional leadership is related to teacher competency. Instructional leaders provide the support teachers still need as they learn. As these teachers improve, instructional leaders become more effective in their schools. The findings benefit education decision-makers and school administrators, particularly for the Education Strategic Plan 2024-2028. They provide a basis for evidence-based adjustments in teacher training. Further studies are needed to explore these dynamics in different contexts to support the national education system.

7. Recommendation

The study's findings on the relationship between instructional leadership and teacher competency lead to the following recommendations: The BTEC administration should codify management training programs to ensure leadership principles are used consistently to support pre-service teachers. Leadership has to do with competence. Thus, leadership must ensure that pre-service teachers receive feedback they can use to grow. This feedback should happen often. The good thing about "learning" is that it really works. We can see this from what people said in the findings. So, the institution should make "learning" a big part of the teacher education curriculum. This will help teachers improve their practice. Then we can give them the help they need, rather than just giving everyone the same thing. This way, we can ensure our support makes a real difference for the teachers who need it most. Leadership and competency are important; we should focus on them.

Author Contributions: All authors contributed to this research.

Funding: This study received no funding.

Conflict of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics Approval: Not applicable.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Darling-Hammond, L. (2006). *Powerful Teacher Education: Lessons from Exemplary Programs*. San Francisco: John Wiley & Sons, Inc
- Darling-Hammond, L., Burns, D., Campbell, C., Goodwin, A. L., Hammerness, K., Low, E. L., & Zeichner, K. M. (2017). *Empowered Educators: How High-Performing Systems Shape Teaching Quality around the World*. John Wiley & Sons.
- Dismore, H., Campbell-Barr, V., Manning, R., & Warwick, P. (2024). A relational approach to knowledge exchange in higher education. *Studies in Higher Education*, 49(12), 2534–2545. <https://doi.org/10.1080/03075079.2024.2312404>
- Hallinger, P. (2011). Leadership for Learning: Lessons from 40 Years of Empirical Research. *Journal of Educational Administration*, 49, 125-142. <https://doi.org/10.1108/09578231111116699>
- Khan, S. (2018). *The teacher competency framework: Elevating pedagogical standards in developing education systems* (Report No. 45). UNESCO.
- Stronge, J. H. (2018). *Qualities of Effective Teachers*. ASCD.

Index**Correlations**

		IL	TC
IL	Pearson Correlation	1	.538**
	Sig. (2-tailed)		.000
	N	303	303
TC	Pearson Correlation	.538**	1
	Sig. (2-tailed)	.000	
	N	303	303

** . Correlation is significant at the 0.01 level (2-tailed).

The Degree of Practice of Islamic Education Teachers in Karak Governorate for Their Roles in Light of Cognitive Load Theory

Abdalkader Saleh Abdalkader Al-Hjooj¹

¹ The World Islamic Sciences & Education University — W.I.S.E, Faculty of Educational Sciences, Department of Curriculum and Instruction. Email: abdalkader.alhjooj@wise.edu.jo
ORCID ID: <https://orcid.org/0009-0006-0553-9170>

Abstract

This study aimed to examine the extent to which Islamic Education teachers in Al-Karak Governorate perform their instructional roles in accordance with Cognitive Load Theory. It also explored whether these practices vary according to teachers' gender, academic qualification, and years of teaching experience. The study employed a descriptive–analytical research design and utilized a questionnaire developed by the researcher as the primary data collection instrument. The questionnaire comprised 30 items organized into three dimensions: intrinsic cognitive load, extraneous cognitive load, and germane cognitive load. Its validity and reliability were established prior to administration. A simple random sample of 235 male and female teachers of Islamic Education in Al-Karak Governorate during the 2025–2026 academic year participated in the study. The results revealed that teachers' perceived level of practicing their instructional roles in light of Cognitive Load Theory was moderate. Of the three dimensions, intrinsic cognitive load ranked highest, followed by extraneous load, while germane load ranked lowest. Furthermore, the results revealed no statistically significant differences in teachers' implementation of their instructional roles based on gender, academic qualification, or years of teaching experience. In light of these findings, the study proposed several recommendations. Most notably, it emphasized the need for the Ministry of Education to provide specialized professional development programs that equip teachers of Islamic Education with strategies grounded in Cognitive Load Theory. The study also recommended conducting further research to examine the effectiveness of curriculum units designed according to Cognitive Load Theory and their impact on students' academic achievement and critical thinking skills in Islamic Education.

Keywords: Degree of Practice, Teachers of Islamic Education, Cognitive Load Theory

1. Introduction

The rapid expansion of knowledge represents one of the foremost challenges to education in the twenty-first century, placing increasing demands on learners' working memory. Consequently, the theory of cognitive load has emerged as a theoretical foundation for understanding how the human mind processes and stores information efficiently. It asserts that effective learning occurs when instructional materials are designed to balance three types of cognitive load: **intrinsic** load, which relates to the complexity of the material; **extrinsic (extraneous)** load,

which relates to poor material design; and cognitive (**germane**) load associated with meaningful learning that builds cognitive schemas (Al-Atoum, 2004).

This theory is of great importance in all academic subjects, especially in Islamic Education, which is characterized by the depth and complexity of its content—ranging from the interpretation of the Holy Quran and the Prophetic Sunnah, to lessons on Islamic creed, including the Islamic conception of the universe, humanity, and life, to the legal rulings concerning the actions of accountable individuals, and extending to the derivation and practical application of moral values. Presenting this intricate content in a traditional manner may accumulate external cognitive load, hinder deep understanding, confine learning to rote memorization and recitation, and prevent the development of thinking skills; thereby widening the gap between theory and reality (Al-Fayl, 2015). Effective teaching extends beyond the mere transmission of information; it aims to achieve meaningful learning outcomes and foster desired changes in students' knowledge and behavior; This has added new roles to the Islamic education teacher, enabling them to create an interactive learning environment in the classroom to effectively bring about the desired changes in student behavior (Al-Dihyan, 2001).

2. Statement of the Problem

Cognitive load constitutes a major barrier to achieving intended educational outcomes in various study subjects. (Fadel, 2014) attributed the reason for this to teaching being based on traditional teaching methods that continuously present information to students, with the students' role limited to listening and memorizing the information presented to them, neglecting their thinking skills; this may hinder the understanding, encoding, processing, and storage of information in working memory. This was confirmed by the study of Hammad (2004), which found that the teaching methods used by Islamic education teachers cause weakness in students' interaction with their teachers, and at the same time do not keep pace with what modern trends in teaching methods recommend. In contrast, the results of the Tracy (2004) study showed that time pressure, stress, and the abundance of information presented in the textbook led to the occurrence of cognitive load, while the study of Si & Kim (2011) presented a solution to this through presenting instructional materials sequentially from simple to complex, which leads to reducing cognitive load among students.

The discrepancy between the complexity of Islamic Education content and students' cognitive processing capacity may exceed learners' abilities if not managed effectively. This is especially true if teachers do not do their jobs in a way that helps students learn without getting too confused. So the issue with this study is that we do not really know how well teachers of education in Jordan are doing their jobs when it comes to helping students learn. We also need information about this topic, especially in the Karak Governorate area. To understand the issue, we need to answer some questions about education and how teachers are doing their jobs in relation to how students learn⁶

1. To what extent do Islamic Education teachers in Al-Karak Governorate practice their instructional roles in accordance with Cognitive Load Theory?
2. Are there any statistically significant differences ($\alpha \leq 0.05$) in the extent to which Islamic Education teachers practice their instructional roles in accordance with Cognitive Load Theory based on gender, educational qualification, and years of teaching experience?

3. Objectives of the Study

The primary objectives of this study are:

1. To assess the degree to which Islamic Education teachers in Al-Karak Governorate apply Cognitive Load Theory within their instructional practices.
2. To determine whether significant differences exist in teachers' self-perceptions of their practice based on demographic variables such as gender, experience, and educational background.

4. Significance of the Study

The significance of the study lies in two aspects: theoretical and applied, which serve both the scientific and practical fields for Islamic education teachers:

Theoretical Significance: The study may contribute to providing a theoretical framework on cognitive load theory and how to employ it in Islamic education, which is considered an important approach that contributes to developing the learner's personality in a balanced manner and developing his thinking skills through applying the principles of the theory in a way that achieves positive results in learning, and contributes to developing students' thinking and learning by defining teachers' roles in light of this theory.

Applied Significance: This study may provide Islamic education teachers with a practical work framework and clear practical steps to redesign their lessons and activities in a way that reduces distraction and focuses on deep understanding according to cognitive load theory, by clarifying the roles of Teachers in the field of Islamic Education in light of the theory. The findings are also expected to yield actionable recommendations for curriculum designers and practitioners in Islamic Education so that the principles of cognitive load management are adopted when designing textbooks and their accompanying instructional aids.

5. Terms of the Study

Study Terminology: This study uses the following terms:

- Islamic Education Teacher: A teacher appointed by the Ministry of Education and tasked with teaching Islamic Education in schools within the Karak Governorate for the 2025–2026 school year.
- Cognitive load: The amount of mental effort expended by a learner's working memory to process new information; this effort is measured based on the limited capacity of working memory (Sweller, 1980).
- Teachers' Roles in Cognitive Load: The instructional strategies employed by Islamic education teachers to address extrinsic load, control intrinsic load, and manage learning-related load in order to direct students' mental effort toward understanding and effective application (Van Merriënboer, J. J. G., & Sweller, J., 2005). These can be defined procedurally as the roles related to managing intrinsic load, managing extrinsic load, and managing learning-related load performed by Islamic education teachers, which are measured using a tool designed for this purpose.

6. Limits and Delimitations of the Study

Human Limitations:

The study is limited to a sample of male and female Teachers in the field of Islamic Education in schools affiliated with Karak Governorate.

Temporal Limitations:

The study was conducted during the second semester of the academic year (2025/2026).

Spatial Limitations:

The participants in this study consisted exclusively of male and female Islamic Education instructors in Al-Karak Governorate.

Delimitations of the Study:

The generalization of the study results is determined by the reliability and validity of the study instrument and by the extent to which the study sample responded accurately to the instrument items.

7. Theoretical Framework

Cognitive load theory was developed by John Sweller and his colleagues, and sought to explain how memory works during learning, and the teacher's role in effectively organizing students' learning with a good impact. This theory refers to the total amount of mental activity in working memory at a given time, and the number of elements that must be attended to; this means that there is an upper limit of cognitive elements that memory cannot exceed at once. When the cognitive elements to be remembered are between (5-8) elements, the task is recallable if the learner exerts some effort, but when their number reaches more than (10) elements, the task becomes difficult for

most learners, and distracts the learner's mind, especially if the way the instructional material was presented was inappropriate, and causes difficulty in the material; and this forms what is called cognitive load (Abu Riyash, 2007).

Cognitive load refers to the set of mental activities that occupy the capacity of working memory during a certain time (Sweller, 2008), and (Cooper, Ford, Chandler, Sweller, 1998) added that it is the total amount of mental activity in working memory over a certain time, measured by the number of cognitive units or elements, and it is divided into three overlapping types, referred to by (Sweller, 2008), which can be projected onto the Islamic education subject, namely:

Intrinsic Load: The burden resulting from the difficulty and complexity of the content and the degree of interconnection of its elements in Islamic education, and it is related to the extent of the difficulty of jurisprudential or creedal concepts, their complexity, and the interconnection of their elements.

Extraneous/External Load: The burden resulting from poor design of materials or the use of inappropriate teaching strategies that do not achieve learning outcomes and lead to student distraction, such as unnecessary repetition or unclear instructions.

Load Related to Internal (Germane) Learning: The organized burden of exploiting students' abilities in designing and building cognitive schemas in long-term memory, thereby achieving deep understanding with the possibility of applying it.

Cognitive load theory divides knowledge into primary and secondary. Primary knowledge refers to knowledge that is learned and acquired through social interaction and the reality into which the individual was born, such as speaking the mother tongue, routine social relationships, and the skills that individuals use in solving problems they face. As for secondary knowledge, it includes: knowledge that the individual acquires through exerting significant mental and cognitive effort, and it is intentional knowledge, and is the focus of cognitive load theory, which sought to establish a mechanism for processing information according to principles on which this theory is based, represented in information storage, borrowing and cognitive reorganization, incidental (creative) production, narrow limits of change, and linking and organizing the environment (Sweller & Sweller, 2006).

The huge amount of information that students receive may harm their working memory, impair its functioning, and weaken its ability to perform its required role, and this may make students feel failure and lower their self-confidence, due to the abundance and overlap of information that requires students to attend to and process it within a specific time, as working memory bears the burden of processing large and complex information at once, which limits its ability to do so, and imposes on it a cognitive load that affects the effectiveness of learning. Also, not giving the learner enough time to think about the material presented to him limits his ability to employ that information in solving the problems and situations he encounters. Here appears the importance of cognitive load theory, which uses scientific methods to reduce the level of cognitive load on working memory, and increases the effectiveness of the teaching and learning process, leading students towards theoretical thinking and creativity, and applied grounding in various fields of science and knowledge (Abdel Hadi, 2010).

Cognitive load theory is based on rules that reduce the cognitive load on the learner through analyzing information and defining parts, using interconnected presentations, deleting repeated information, organized exploration of the issues or questions presented instead of repetition, presenting effects and the audio narrative simultaneously rather than sequentially, and presenting worked examples as alternatives to the conventional problem agreed upon in existing teaching (Abu Riyash, 2007).

Cognitive load theory also sought to develop the learning process by presenting an organized and effective instructional design that does not lead to an overload on working memory during the individual's learning process, and providing curriculum designers with solid foundations that enable them to build the curriculum according to scientific bases, and working to reduce extraneous cognitive load to save learning time and mental effort in the

learning process instead of depleting his abilities, and paying attention to repeated training to increase his ability to understand the content (Kalyuga, 2006).

What distinguishes cognitive load theory is that it improves teachers' performance, because it drives them to draw various pictorial and semantic cognitive schemas and different meanings and their uses in classroom learning situations, and to use strategies that link knowledge together with relationships that contribute to transferring useful experiences to learners (Qatami, 2013).

The Islamic education teacher can reduce the extraneous load resulting from a defect in material design by applying several principles, including the duality principle and temporal and spatial approximation, so the teacher avoids presenting the same information in two different forms at the same time, such as presenting words associated with the picture simultaneously rather than sequentially to ensure their joint processing in working memory. Also, the coherence principle by removing any unnecessary or distracting information, graphics, or attractive images, and the principle of fully worked examples for new learning instead of having students solve problems, such as presenting fully solved jurisprudential problems, and giving them similar problems to enable them to build cognitive schemas and reduce cognitive load (Al-Otaibi & Mustafa, 2017).

As for managing intrinsic load, the teacher can reduce the load resulting from content complexity through the segmentation principle by dividing the lesson into interconnected topics presented sequentially, such as dividing the prayer lesson into sequential topics, providing appropriate time to understand the sub-topic before moving to the next topic. The teacher also relies on the pre-training/advance organizer principle, so that the transition is from more general concepts to less general concepts to achieve an organized hierarchical structure of concepts, and then moving to the details of sub-rulings, such as the conditions of validity of prayer and its pillars, to link new knowledge with old knowledge in a coherent manner that achieves an organized hierarchical structure of information. The teacher also activates guided and deep discussion by asking interpretive questions that require the learner to think and analyze, such as: " Why did Islam prohibit cheating? And how can we avoid it? Instead of the teacher being content with asking the learner what the ruling is on cheating in Islam (Abu Riyash, 2007).

The teacher can reduce the load related to learning (germane load) by gradually moving from presenting fully worked examples to the learner solving problems himself, and also activating prior schemas, through prior comparison, summarization, or guided discussion that leads the learner to link new information with old information, to achieve the construction process in the learner's long-term memory (Abu Ghouri & Sheldan, 2025). This can be done through solving jurisprudential problems related, for example, to analogy (qiyas), which is one of the sources of Islamic legislation; for example, if the teacher gives a worked example about wine in terms of its ruling and the reason for its prohibition, he directly assigns students to solve similar problems such as drugs, so that the learner reaches its ruling, and knows that the reason for the prohibition of drugs is the same as the reason for the prohibition of wine, in that both impair the mind.

The teacher can work on adapting cognitive load when assessing and following up on students through tests that lead the learner to apply the Sharia ruling to cases he has not previously dealt with, and providing organized feedback so that the notes are specific and focused on the errors that prevent the learner from building the hierarchical schema. To ensure success, the teacher must conduct short diagnostic tests at the beginning of the topic to determine the learner's prior knowledge; if it is weak, the teacher reduces the complexity of the subsequent topic to lower intrinsic load.

8. Previous Studies

In light of the scarcity of previous studies dealing with the topic of the study, several previous studies related to cognitive load theory, which are relevant to the topic of the current study, were reviewed, and the following is a presentation of these studies.

Al-Zahrani (2024) conducted a study that aimed to identify the extent of the effectiveness of using methods of reducing cognitive load in teaching at the childhood stage. The study used the descriptive analytical approach; the

study sample, selected by the simple random method, consisted of 100 kindergarten teachers in the schools affiliated with the cities of Abha and Khamis Mushait. The results found the effectiveness of using methods of reducing cognitive load in achieving positive learning outcomes, such as practicing drawing cognitive schemas and linking knowledge with relationships, to transfer experiences to children, and recommended the necessity of adopting the principles, methods, and strategies of reducing cognitive load in teaching the kindergarten curriculum. Al-Otaibi and Mustafa did a study in 2017. They wanted to see if using a formality strategy based on cognitive load theory would help year secondary female students in Saudi Arabia do better in the jurisprudence subject. The study used an approach called quasi-experimental. What they found out was that using a formality strategy based on cognitive load theory really helped year secondary female students do better in the jurisprudence subject. Al-Otaibi and Mustafa found that this strategy was effective for these students. The students were in Saudi Arabia. The subject was jurisprudence. Al-Otaibi and Mustafa looked at how the students did, after using the formality strategy based on cognitive load theory. The result was that the students did better in jurisprudence.

Bishay's (2016) study aimed to investigate the effectiveness of using instructional scaffolding in developing engineering problem-solving skills and reducing cognitive load among a group consisting of 62) students of second-year preparatory students in Assiut city, divided into two groups: one experimental and the other control. After conducting the research experiment, a test of engineering problem-solving skills was applied, and the NASA cognitive load scale was used. The research results revealed a statistically significant effect of using instructional scaffolding in developing engineering problem-solving skills and reducing cognitive load.

Al-Harithi's (2014) study investigated the relationship of cognitive load with perceptual skills represented in auditory perception, visual perception, mental perception, and motor perception, among sixth-grade primary students, and the sample consisted of 120 sixth-grade students with academic learning difficulties in various schools in the Riyadh region. The study found a statistically significant negative relationship between students' scores on the cognitive load scale and their scores on the various perceptual skills scale, and recommended the necessity of reducing cognitive load for students with learning difficulties or attempting to provide solutions to develop variables related to cognitive load, such as working memory and long-term memory.

Amadiou, Marine & Laimay (2010) conducted a study that aimed to reveal the effect of directing attention to high-content and low-interactivity elements of instructional material in reducing cognitive load during learning. The study used the quasi-experimental approach, and the sample consisted of 68 male and female students from the psychology department at the University of Toulouse in France. Participants were divided into two experimental groups: one group was presented with a surgical procedure without directing students' attention to the elements of the instructional material, while the second experimental group was shown the surgical operation with directing students' attention to the elements related to the instructional material. The study used the Paas scale to measure cognitive load, and the results revealed a decrease in the level of cognitive load when directing students' attention towards interactive elements, and an increase in cognitive load when students' attention was not directed to the elements of the instructional material.

Tracy did a study in 2004 to see how much load primary stage students had. Tracy used a way of studying that looks at things and tries to understand them, and Tracy picked one hundred students, boys and girls, to be in the study. The study by Tracy found that there are a lot of things outside of school that affect a student's load, like the science books and materials that the student can use, how tired the student is, and how long it takes to finish the work. These things, like the science materials and the time to finish the work, and the tiredness, play a part in making cognitive load higher or lower for students. For the students in the study, according to Tracy, the cognitive load is affected by these things.

Abu Jouda (2004) conducted a study that investigated the effect of using a teaching-learning program derived from cognitive load theory in developing critical thinking skills among a sample of tenth-grade basic students in Zarqa Governorate in Jordan. To achieve the study's objective, the quasi-experimental approach was used, and 88 male and female students were selected from schools in Zarqa Governorate, where a teaching-learning program based on cognitive load theory was built. The results showed the effectiveness of the teaching-learning program derived

from cognitive load theory in developing critical thinking skills among a sample of tenth-grade basic students in Zarqa Governorate in Jordan.

9. Comments on Previous Studies

Previous studies varied; some of them studied the effect of modern programs and strategies on reducing cognitive load, using the quasi-experimental approach, such as the study of Al-Otaibi & Mustafa (2017), the study of Abu Jouda (2004), and the study of Bishay (2016). Others sought to measure the level of cognitive load among students using the descriptive analytical approach, such as the study of Al-Zahrani (2024) and the study of Amadiou, Marine & Laimay (2010). Others investigated the relationship between cognitive load and other variables using the descriptive correlational approach, such as the study of Al-Harithi (2014).

Regarding the results reached by previous studies, some of them showed the effectiveness of modern strategies in reducing cognitive load, such as the study of (Jaber, 2016) and the study of (Amadiou, Marine & Laimay, 2010), and others showed the effect of some variables on increasing cognitive load, such as the study of (Tracy, 2004), which showed that the scientific material available to the student, mental fatigue, and the time taken to complete the task are among the factors that control the level of cognitive load, and the study of (Al-Zahrani, 2024), which showed the effectiveness of using methods of reducing cognitive load in teaching at the childhood stage.

The study benefited from these studies in enriching the scientific material of the current study, designing the study instrument, and drawing on the results of a number of these studies to formulate the problem of the current study, while the current study was distinguished from previous studies in that it investigated the degree of practice of Islamic education teachers in Karak Governorate for their roles in light of cognitive load theory, as no study addressing the topic of the study was found — to the extent of the researcher's review of the literature.

10. Methodology and Design

Study Methodology: The study adopted the descriptive analytical approach for its suitability to the nature of the study.

Study Population: The study population consisted of all Islamic education teachers in Karak Governorate for the academic year (2025/2026), numbering 756 male and female teachers, according to the statistics of the directorates of education in the governorate.

Study Sample: The study sample, selected by the simple random method, consisted of 235 participants, where 250 questionnaires were distributed, 242 of which were retrieved, and after entering the data, 7 questionnaires were excluded for being unsuitable for analysis. Thus, the sample subject to analysis consisted of (235) male and female teachers, at a rate of (31.08%) of the study population, and Table 1 shows the characteristics of the study sample according to the variables of gender, academic qualification, and years of experience.

Table 1: Distribution of study sample members according to their variables

Variable	Category	N	Percentage (%)
Gender	Male	108	46.0%
	Female	127	54.0%
	Total	235	100.0%
Academic Qualification	Bachelor's	161	68.5%
	Postgraduate Studies	74	31.5%
	Total	235	100.0%
Years of Experience	5 years or less	49	20.9%
	6-10 years	125	53.2%
	More than 10 years	61	26.0%
	Total	235	100.0%

Study Instrument: To achieve the objective of the study, the study instrument was developed based on the relevant theoretical literature and previous studies, such as the study of (Abu Ghouri & Sheldan, 2025) and the study of (Al-Otaibi & Mustafa, 2017), where the instrument in its initial form consisted of a questionnaire containing (30) items on the degree of practice of Islamic education teachers for their roles in light of cognitive load theory.

11. Validity of the Instrument

The validity of the instrument was verified by two methods, namely:

First: Validity of the arbitrators (jury validity): The questionnaire was presented to (10) specialized arbitrators in the specializations of curriculum and instruction, and supervisors of the Islamic education subject, to verify its ability to achieve its intended purpose, and to ensure the clarity and soundness of the wording of the items and their validity for measuring what they were designed to measure, and to make any modification of deletion, addition, or rewording of items and their suitability to the topic. Based on the arbitrators' modifications to the items of the study instrument, with an agreement rate of (80%) of the arbitrators, this was considered an indicator of the validity of the items. The arbitrators' observations were taken into account. The required and appropriate modifications were made, and the questionnaire settled in its final form on (30) items, distributed across three domains, namely: intrinsic load, extraneous load, and germane load.

Second: Internal Consistency Validity

Internal consistency was assessed to verify the reliability of the research instrument. To do this, we used a method that checks how well the different parts of the instrument work together. This method is part of making sure the instrument is really measuring what it is supposed to measure. We looked at how well each question on the questionnaire matched up with the rest of the questions in that section and with the score. We tried this out with a group of 30 teachers, both men and women, from the same group of people we were studying, and from outside of that group. The results are shown in Table 3)

Table 3: Correlation coefficients of items with the domain and the total score of the questionnaire

Item No	Correlation with Domain	Correlation with Total	Item No	Correlation with Domain	Correlation with Total
First Domain: Managing Intrinsic Load					0.668**
1	0.483**	0.327*	6	0.673**	0.456**
2	0.727**	0.556**	7	0.653**	0.444**
3	0.587**	0.624*	8	0.575**	0.504**
4	0.642**	0.317*	9	0.624**	0.509**
5	0.894**	0.728**	10	0.929**	0.757**
Second Domain: Extraneous Load					0.796**
11	0.910**	0.724**	16	0.832**	0.747**
12	0.814**	0.716**	17	0.889**	0.732**
13	0.823**	0.710**	18	0.737**	0.767**
14	0.888**	0.675**	19	0.849**	0.653**
15	0.883**	0.695**	20	0.898**	0.753**
Third Domain: Germane Load					0.665**
21	0.871**	0.590**	26	0.831**	0.467**
22	0.804**	0.552**	27	0.865**	0.571**
23	0.741**	0.664**	28	0.825**	0.654**
24	0.742**	0.657**	29	0.775**	0.631**
25	0.669**	0.632**	30	0.564**	0.769**

*The results are important at the level we care about *they are really important at a strict level.

Table 3 shows that all the connections between the study tool and the category, as well as with the total score of the tool, were important at the level we care about, where the connections between the items and the total score of the questionnaire were between 0.317 and 0.769. The connections between the items and the category they belong

to were between 0.456 and 0.910. The connections between the categories and the total score of the tool were between 0.665 and 0.796. This shows that the questionnaire is a measure of what it is supposed to measure, and it makes the results of the questionnaire more reliable. The questionnaire is a tool because all the questionnaire parts are connected to the total score of the questionnaire and to the category of the questionnaire, which is what we want to see in a good study tool, like the questionnaire.

12. Reliability of the Study Instrument

To see if the study tool was good, we tried it out with a group of thirty teachers, both men and women, from the main group of people we were studying and from outside of it. We wanted to make sure it was reliable. We used something called the Cronbach's alpha equation to check this. The results are shown in Table 2. This table shows how reliable the study tool is for each part and for the thing. We looked at the study tool to see if it was working well for each domain and for the study tool as a whole. The study tool was checked to make sure it was giving us information.

Table 2: Reliability coefficient values for each domain and for the instrument as a whole using the Cronbach's alpha equation

Domain	Cronbach's Alpha Coefficient
Managing Intrinsic Load	86%
Managing Extraneous Load	82%
Enhancing Germane Load	83%
Total	87%

The results in Table 2 show that the reliability of the instrument is pretty good. We used the Cronbach's alpha method to check this. The results for each part of the instrument ranged from 82% to 86%. For the instrument, the result was 87%. The researcher Odeh said in 2014 that these numbers are good enough for what we're trying to do in this study.

13. Scoring of the Study Instrument

The five-point Likert scale was adopted for scoring the instrument, giving each item one degree out of its five degrees (I practice it to a very great degree, I practice it to a great degree, I practice it to a moderate degree, I practice it to a small degree, I do not practice it at all), representing numerically (5, 4, 3, 2, 1) respectively, and the following criterion was adopted in judging Islamic education teachers' estimates of their degree of practice of cognitive load management strategies (Odeh, 2014):

Criterion for judging arithmetic means:

1.00–2.33: Low degree

2.34–3.67: Moderate degree

3.68–5.00: High degree

I came up with this estimate by taking the score, which is 5, and breaking it down into three parts that are all the same size. These parts are based on a scale that goes from 1 to 5. To figure this out, I used a math problem. I took the answer choice from the study and subtracted the lowest answer choice from the study. Then I divided that number by the number of categories, which are high, moderate, and low. This is similar to what Odeh did in 2014:

$$\frac{N-1}{3} = \frac{5-1}{3} = 1.33$$

14. Statistical Treatment

The Statistical Package for the Social Sciences, which we call SPSS, helped us find answers to our study questions using statistical methods. The following outlines the procedure used in this study:

To answer the question, we calculated the average and standard deviation.

For the question, we used a three-way multivariate analysis of variance, or MANOVA for short, and the Statistical Package for the Social Sciences, which is SPSS.

Presentation and Discussion of Results

Results Related to the First Question

What is the degree of practice of Islamic education teachers in their roles in light of cognitive load theory?

To answer the question, arithmetic means and standard deviations were calculated for Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory, and Table 3 shows that.

Table 3: Arithmetic means and standard deviations of Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory, ranked in descending order

No	Domain	Mean	SD	Rank	Practice Degree
1	Intrinsic Load	3.65	0.89	1	Moderate
2	Extraneous Load	3.63	0.91	2	Moderate
3	Germane Load	3.56	0.81	3	Moderate
Total Degree		3.61	0.83		Moderate

Table 5 shows that teachers' ratings of the extent to which they practiced their roles in accordance with Cognitive Load Theory were moderate across all domains, with an overall mean score of 3.61 and a standard deviation of 0.83. Among the three domains, intrinsic load ranked first, with a mean of 3.65 and a standard deviation of 0.89. This was followed by extraneous load, which ranked second, with a mean of 3.63 and a standard deviation of 0.91. Germane load ranked third and last, with a mean of 3.56 and a standard deviation of 0.81.

The moderateness of the degree of teachers' practice of their roles in light of cognitive load theory, and it not being low, can be attributed to teachers' good awareness of the importance of cognitive load management, in addition to the fact that Islamic education teachers automatically practice some cognitive load reduction strategies, such as simplifying the concept, segmenting the topic, and using audio-visual and technological aids. The result may also be attributed to teachers' professional intuitive awareness that presenting information in a congested manner may distract students; therefore, they seek to control their teaching to ensure learners' comprehension of Islamic education topics. Therefore, the results indicated that their roles in the domain of intrinsic load related to the nature of the material came in the highest ranks of practice, as the teacher focuses on the core of the educational process by presenting the material in a way that suits students' abilities, and this result is also related to teachers' serious attempt to adapt to the developed curricula in recent years, which have become more focused on thinking skills instead of memorization, which contributed to their practice of roles that reduce distracting and unnecessary cognitive load.

However, the moderateness of this practice and the failure of its level to rise indicates the presence of educational or cognitive obstacles that contributed to that, including the absence of deep specialized training; cognitive load theory is a precise theory that requires an understanding of the three types of load (intrinsic, extraneous, germane), and teachers may lack such precise training, as the teacher practices according to what he knows only without being precisely familiar with his roles in light of load theory, especially the extraneous type related to content design. The result can also be explained by the abundance of tasks assigned to the teacher, especially those not related to teaching the subject, such as duty shifts, supervision periods, and school-related activities, and assigning teachers administrative tasks that limit their ability to practice their roles. The researcher believes that teachers may succeed in simplifying the lesson to reduce cognitive load, but they may face difficulty in moving to the more important step, which is motivating students to build knowledge deeply, which is a matter that requires possessing and practicing skills that go beyond classroom management limits, and requires intensive training to ensure they carry out their roles efficiently.

First Domain: Managing Intrinsic Load

Table 4: Arithmetic means and standard deviations of Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory in the domain of (intrinsic load), ranked in descending order

No	Item	Mean	SD	Rank	Practice Degree
9	I balance between the theoretical and practical aspects in presenting lessons	3.85	1.06	1	High
8	I select from the Sharia evidence what serves the lesson's objective without burdening the student's mind	3.76	1.19	2	High
2	I present religious concepts from simple to compound in a way appropriate to the age stage	3.67	1.12	3	Moderate
3	I use prophetic similes to bring unseen (metaphysical) concepts closer to students' minds	3.67	1.06	3	Moderate
7	I check prior experiences before starting to explain a new jurisprudential rule	3.66	1.20	4	Moderate
1	I divide long jurisprudential lessons into main elements to facilitate comprehension of Sharia rulings	3.64	1.03	5	Moderate
5	I link fundamentals (usul) and branches (furu') to reduce students' distraction	3.60	1.21	6	Moderate
10	I work on highlighting the wisdom behind Sharia rulings to facilitate their understanding and memorization	3.56	1.30	7	Moderate
4	I use lesson outlines to control the logical sequence in presenting the Holy Verses	3.55	1.13	8	Moderate
6	I use easy language to explain difficult foundational (usuli) and purposive (maqasidi) terms	3.53	1.16	9	Moderate
Overall Mean		3.65	0.89		Moderate

The arithmetic mean of the items of the (intrinsic load) domain ranged between (3.53-3.85), and came with degrees ranging between high and moderate, and item No. (9), whose text is (I balance between the theoretical and practical aspects in presenting lessons), came in first place with an arithmetic mean of (3.85) and a standard deviation of (1.06), which can be attributed to the fact that Islamic education lessons rely on both the theoretical and practical aspects, and achieving sound understanding requires balancing between the two aspects, such as teaching jurisprudence and personal status topics, for example.

Item No. (6), whose text is (I use easy language to explain difficult foundational (usuli) and purposive (maqasidi) terms), came in the last place, with an arithmetic mean of (3.53) and a standard deviation of (1.16), and the delay of this item behind the rest of the items may be attributed to the fact that some creedal terms may be abstract and need to be simplified by linking them to the reality experienced by students and linking them to tangible real-life examples, but some teachers may lack the experience to do so.

Second Domain: Extraneous Load

Table 5: Arithmetic means and standard deviations of Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory in the domain of (extraneous load), ranked in descending order

No	Item	Mean	SD	Rank	Practice Degree
18	I design worksheets characterized by clarity and that do not require mental effort to achieve the goal	3.75	1.04	1	High
11	I organize the board/presentation in a way that prevents congestion of information	3.73	1.11	2	High
16	I avoid sudden transitions between different Sharia topics by organizing class time	3.70	1.03	3	High

19	I choose instructional aids that contribute to an accurate understanding of topics	3.68	1.08	4	High
17	I avoid unnecessary verbal repetition that weakens students' concentration	3.64	1.16	5	Moderate
13	I control the pace of explanation so that the amount of information presented suits the student's ability	3.63	1.06	6	Moderate
12	I use visual media related to Islamic education topics, especially the Prophetic biography (Seerah)	3.59	1.29	7	Moderate
14	I avoid distracting the student's mind with examples unrelated to the lesson topic	3.57	1.14	8	Moderate
20	I manage side conversations among students to ensure their discussion stays within the lesson topic	3.54	1.24	9	Moderate
15	I take care to create an appropriate classroom environment that enables the learner to reflect and think	3.52	1.11	10	Moderate
Overall Mean		3.63	0.91		Moderate

The arithmetic mean of the items of the (extraneous load) domain ranged between (3.52-3.75), and came with degrees ranging between high and moderate, and item No. (18), whose text is (I design worksheets characterized by clarity and that do not require mental effort to achieve the goal), came in first place with an arithmetic mean of (3.75) and a standard deviation of (1.04), which may be attributed to the role of properly prepared worksheets in achieving deep understanding and consolidating information among students, as it is also a formative assessment tool that enables the teacher to measure students' level of comprehension of topics, and identify learning gaps to address them.

Item No. (15), whose text is (I take care to create an appropriate classroom environment that enables the learner to reflect and think), came in the last place, with an arithmetic mean of (3.52) and a standard deviation of (1.11), and its delay may be attributed to the fact that the conditions of some classrooms may limit the teacher's ability to provide an appropriate classroom environment, such as the large number of students or their crowding in classrooms, or the unavailability of appropriate technological aids for use in topics that require mental effort and concentration from the learner.

Third Domain: Germane Load

Table 6: Arithmetic means and standard deviations of Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory in the domain of germane load, ranked in descending order

No	Item	Mean	SD	Rank	Practice Degree
29	I link jurisprudence topics to students' reality to facilitate building cognitive schemas	3.75	1.03	1	High
22	I make students extract lessons learned from Islamic education topics in their own style	3.72	1.06	2	High
21	I motivate students to link the moral values derived from Islamic education with their practiced behavior	3.70	1.08	3	High
30	I encourage students to independently derive Sharia rulings after presenting the appropriate data	3.66	1.05	4	Moderate
24	I employ critical thinking to compare jurisprudential rulings and their applications in the current era	3.63	1.07	5	Moderate
26	I link creed topics with reflecting on God's creation in the universe to enhance students' level of faith	3.61	1.07	6	Moderate
27	I pose questions that provoke thinking about the objectives (maqasid) of Islamic law	3.56	1.12	7	Moderate
28	I direct students to rephrase the Sharia evidence in a sound linguistic style	3.54	1.23	8	Moderate

25	I encourage students to mentally summarize what has been memorized of Quranic verses and prophetic hadiths	3.26	0.66	9	Moderate
23	I use concept maps to link Islamic education topics, especially jurisprudential ones, with each other	3.18	0.77	10	Moderate
Overall Mean		3.56	0.81		Moderate

The arithmetic mean of the items of the (germane load) domain ranged between (3.18-3.75), and came with degrees ranging between high and moderate, and item No. (29), whose text is (I link jurisprudence topics to students' reality to facilitate building cognitive schemas), came in first place with an arithmetic mean of (3.75) and a standard deviation of (1.03), which can be attributed to the fact that linking jurisprudence to the reality of the student's life transforms theoretical information into familiar contexts, arouses the learner's motivation, and encourages him to invest his mental abilities in deep processing operations to build strong and relevant connections between his new and previous experiences; and enhances the germane load that achieves learning and consolidates information in long-term memory.

Item No. (23), whose text is (I use concept maps to link Islamic education topics, especially jurisprudential ones, with each other), came in the last place, with an arithmetic mean of (3.18) and a standard deviation of (0.77), and its delay may be attributed to teachers' lack of experience in building concept maps and making connections between them in an organized manner that achieves deep understanding among students, as building such maps, especially the complex ones, requires training teachers to prepare them in a way that suits the learner and achieves integrated and coherent conceptual construction in the learner's long-term memory.

Findings Related to the Second Research Question

Are there statistically significant differences at the significance level of ($\alpha \leq 0.05$) in Islamic Education teachers' estimates of the extent to which they practice their instructional roles in light of Cognitive Load Theory attributable to gender, years of teaching experience, and educational qualification?

To answer this question, means and standard deviations were calculated for teachers' ratings of the extent to which they practiced their instructional roles in light of Cognitive Load Theory according to the variables of gender, educational qualification, and years of teaching experience. The results are presented in Table (4):

Table 7: Arithmetic means and standard deviations of Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory according to the variables (gender, years of experience, and academic qualification)

Variable	Type	Means and Standard Deviations	Instrument Domains			Total
			Intrinsic Load	Extraneous Load	Germane Load	
Gender	Male	Mean	3.68	3.62	3.59	3.63
		SD	0.93	0.98	0.83	0.88
	Female	Mean	3.63	3.64	3.53	3.60
		SD	0.85	0.86	0.79	0.80
Academic Qualification	Bachelor's	Mean	3.68	3.65	3.59	3.64
		SD	0.89	0.93	0.82	0.84
	Postgraduate Studies	Mean	3.58	3.60	3.50	3.56
		SD	0.87	0.87	0.79	0.81
Years of Experience	5 years or less	Mean	3.90	3.81	3.72	3.81
		SD	0.76	0.77	0.74	0.72
	6-10 years	Mean	3.58	3.56	3.46	3.54
		SD	0.94	0.95	0.83	0.87
	More than 10 years	Mean	3.60	3.64	3.62	3.62
		SD	0.85	0.94	0.80	0.82

The results in Table 4 show that there are some differences in the average scores and how much they vary for Islamic education teachers when they estimate how well they do their jobs, considering how much they have to think about. This is based on things like whether they're a man or a woman, what kind of degree they have, and how long they have been teaching. To see if these differences are really important, we used a kind of statistical test, which is shown in Table 5.

Table 8: Three-way multivariate analysis of variance for the effect of the variables (gender, academic qualification, and years of experience) on Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory

Source of Variance	Dependent Variable	Sum of Squares	df	Mean Square	F Value	Sig.
Gender – Hotelling's Value 0.015	Intrinsic Load	0.127	1	0.127	0.163	0.687
	Extraneous Load	0.016	1	0.016	0.019	0.890
	Germane Load	0.281	1	0.281	0.432	0.512
	Total	0.064	1	0.064	0.093	0.761
Academic Qualification – Hotelling's Value 0.003	Intrinsic Load	0.257	1	0.257	0.329	0.567
	Extraneous Load	0.065	1	0.065	0.077	0.782
	Germane Load	0.160	1	0.160	0.246	0.621
	Total	0.150	1	0.150	0.217	0.642
Years of Experience – Wilks' Lambda Value 0.955	Intrinsic Load	3.608	2	1.804	2.307	0.102
	Extraneous Load	1.940	2	0.970	1.153	0.318
	Germane Load	2.684	2	1.342	2.062	0.130
	Total	2.511	2	1.256	1.818	0.165
Error	Intrinsic Load	179.875	230	0.782		
	Extraneous Load	193.520	230	0.841		
	Germane Load	149.710	230	0.651		
	Total	158.848	230	0.691		
Total	Intrinsic Load	3313.090	235			
	Extraneous Load	3299.850	235			
	Germane Load	3129.780	235			
	Total	3231.369	235			
Corrected Total	Intrinsic Load	184.127	234			
	Extraneous Load	195.651	234			
	Germane Load	152.908	234			
	Total	161.722	234			

The results presented in Table (5) indicate the absence of statistically significant differences at the significance level ($\alpha \leq 0.05$) in Islamic education teachers' estimates of their degree of practice of their roles in light of cognitive load theory attributable to the variables (gender, academic qualification, and years of experience), and this result can be attributed to the fact that teachers practice similar roles, as the curricula are unified, the work environment is similar, most of them bear administrative and teaching burdens, and they follow specific standards set by the Ministry of Education, in addition to the low level of training received by teachers, regardless of their gender, academic qualification, or years of experience.

15. Recommendations

1. Holding specialized training courses by the Ministry of Education to enable Islamic education teachers to raise their degree of practice of their roles in light of cognitive load theory.
2. Highlighting the role of educational supervisors in enhancing the level of teachers' practice of their roles in light of cognitive load theory through giving applied lessons, and conducting exchange visits that contribute to practically transferring experiences.
3. Holding educational workshops via educational platforms such as Zoom for all Islamic education teachers, aiming to enable them to practice their roles in light of cognitive load theory in a better way.
4. The necessity of training teachers on how to build concept maps and use easy language to explain abstract concepts.
5. Calling on school administrations to the necessity of creating an appropriate classroom environment for learning and providing them with an appropriate atmosphere for thinking and reflection.

16. Suggestions

1. Conducting other studies that investigate the effect of a developed unit in light of cognitive load theory in one of the Islamic education books for the basic stage grades, and measuring its effect on students' achievement and thinking in the Islamic education subject.
2. Conducting other studies that investigate the degree of teachers' practice of their roles in light of cognitive load theory in other specializations and locations.

Funding: This study received no funding.

Conflict of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics Approval: Not applicable.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Abu Jouda, Safia Suleiman. (2004). The effect of a teaching-learning program based on cognitive load theory in developing critical thinking skills, unpublished doctoral dissertation, Amman Arab University, Jordan.
- Abu Riyash, Hussein Mohammad. (2007). Cognitive Learning, 1st ed., Dar Al-Masirah for Publishing, Amman, Jordan.
- Abu Ghouri, Emad Mohammad & Sheldan, Fayez Kamal. (2025). Features of the integrative cognitive curriculum in Islamic education and its contribution to human intellectual and civilizational contribution, *Journal of the Palestinian Educators' League for Arts and Educational and Psychological Studies*, 7(15), 192–226.
- Bishay, Zakaria Jaber Hanawy. (2016). The effectiveness of instructional scaffolding in developing engineering problem-solving skills and reducing cognitive load among second-year preparatory students, *Journal of Mathematics Education*, (8)19, 91–131.
- Al-Harithi, Subhi. (2014). Cognitive load and its relationship with perceptual skills among a sample of sixth-grade primary students with academic learning difficulties, *Educational and Psychological Studies*, Zagazig University, 86(1), 11–48.
- Hammad, Sharif. (2004). Common Islamic education teaching methods used by Islamic education teachers in the upper basic stage in Gaza governorates and justifications for their use, *Journal of the Islamic University of Gaza (Series of Human Studies)*, 12(2), 503–529.
- Al-Dihan, Mohammad Abdulrahman. (2001). The role of primary stage teachers in arousing students' motivation towards learning, *Journal of King Saud University, Educational Sciences and Islamic Studies*, 13(1), 47–85.
- Al-Zahrani, Mona Hashem. (2024). The use of methods of reducing cognitive load in teaching at the childhood stage from the point of view of kindergarten teachers, *Educational and Psychological Studies*, Zagazig University, 137, 1–36.

- Abdel Hadi, Fakhri. (2010). *Cognitive Psychology*, Dar Osama for Publishing, Amman.
- Al-Otoun, Adnan Yousef. (2004). *Cognitive Psychology: Theory and Application*, 1st ed., Dar Al-Masirah, Amman, Jordan.
- Al-Otaibi, Haya & Mustafa, Faten. (2017). The effectiveness of a formality strategy based on cognitive load on the achievement of second-year secondary female students in the jurisprudence subject, *Journal of Scientific Research in Education*, 18(1), 425–443.
- Odeh, Ahmad. (2014). *Principles in Measurement and Evaluation*, Amman: Dar Al-Masirah for Publishing and Distribution.
- Fadel, Nour. (2014). *Cognitive load and its relationship with self-efficacy for confrontation among university students*, published master's thesis, Diyala University, Iraq.
- Al-Feel, Helmy. (2015). *Systemic Intelligence in Cognitive Load Theory*, 1st ed., Anglo-Egyptian Bookshop, Cairo.
- Qatami, Yousef. (2013). *Cognitive Learning and Teaching Strategy*, 1st ed., Amman: Dar Al-Masirah for Publishing and Distribution.
- Amadiou, F., Marine, C., & Laimay, C. (2010). The effects of high prior knowledge on cognitive load during learning with a complex medical simulation. *Computers in Human Behavior*, 27(1), 326–334.
- Cooper, G., Tindall, Ford, S., Chandler, P., & Sweller, J. (1998). Learning by Imaging. *Journal of Experimental Psychology*, 7, 68–82.
- Kalyuga, S. (2006). Assessment of learners' organized knowledge structures in adaptive learning environments. *Applied Cognitive Psychology*, 38, 23–31.
- Si, Jinyun & Kim, Dongsik. (2011). How do instructional sequencing methods affect cognitive load, learning transfer, and learning time? *Educational Research (ISSN: 2141-5161)*, 2(8), 1362–1372, August.
- Sweller, John & Sweller, Susan. (2006). *Natural Information Processing System*, *Evolutionary Psychology*, www.human-nature.com/ep, 4:434–458, Original Article, USA.
- Sweller, John. (2008). *Cognitive Load Theory*, University of New South Wales. WWW.SciTopics.htm
- Sweller, J. (1980). Evolution of human cognitive architecture. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, Vol., San Diego: Academic Press.
- Tracy, Janet Patton. (2004). *Measuring Cognitive Load to Test the Usability of Websites (3rd ed.)*, Indianapolis, IN: Wiley.
- Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17, 147–177.

From the Iceberg Theory to Microcopy: Analyzing the Shift of Minimalist Language from Literature to Digital Products

Hua Ngoc Linh¹, Tran Thanh Du²

^{1,2} Thu Dau Mot University, Ho Chi Minh City, Viet Nam

Correspondence: Tran Thanh Du, Thu Dau Mot University, Ho Chi Minh City, Viet Nam, Tel: 0983947532.
E-mail: dutt@tdmu.edu.vn

Abstract

This research investigates the evolutionary trajectory of minimalist language, tracing its development from 20th-century literary practices to contemporary digital interfaces. It focuses on the adaptation of Ernest Hemingway's "Iceberg Theory" within modern User Experience (UX) writing and microcopy. Employing a mixed-methods approach, this study integrates quantitative corpus linguistics utilizing AntConc software with qualitative functional analysis. A Literary Corpus comprising Hemingway's major works is systematically compared with a UX Writing Corpus drawn from platforms such as Spotify, Slack, Duolingo, and Google. The quantitative findings reveal a Type-Token Ratio (TTR) of 57.99% in UX writing, contrasting with Hemingway's 46.91%. This underscores a significant shift from rhythmic, repetitive prose toward a more functional, action-oriented language. Qualitatively, while both frameworks adhere to Noam Chomsky's (1995) principle of "linguistic economy", their practical objectives differ considerably. Hemingway employs "information gaps" to evoke emotional resonance and foster introspection in readers, whereas UX microcopy emphasizes brevity to minimize cognitive load and facilitate quick user engagement. The study concludes that contemporary minimalism has transformed from a mere aesthetic literary device into a functional "Iceberg Standard" in digital design. This conceptual model effectively conceals complex system logic behind user-friendly interfaces, thereby expanding the realms of applied linguistics and comparative stylistics within human-computer interaction.

Keywords: Comparative Stylistics, Iceberg Theory, UX Writing, Microcopy, Linguistic Economy, Corpus Linguistics.

1. Introduction

1.1 Introduce the Problem

Linguistic minimalism functions not only as an aesthetic trend but also as a highly effective means of communication, particularly as it has evolved significantly from traditional literary forms to contemporary digital formats. Historically, Stylistics has primarily concentrated on analyzing literary works for their aesthetic and emotional impact. As a result, the unique stylistic development of UX writing in the digital age remains underexplored. The fundamental issue lies in the fact that microcopy in digital interfaces is often regarded merely as technical instructions, rather than as a purposeful linguistic structure crafted to influence human behavior.

1.2 Explore the Importance of the Problem

In today's attention economy, where users' cognitive capacity is often limited, simplifying information becomes essential for clear and impactful communication. Despite its morphological brevity, microcopy exerts a disproportionate influence on user behavior and cognitive processing by minimizing friction in interactions and rendering complex technological systems accessible and navigable. Therefore, investigating this shift is crucial to establishing an Iceberg Standard for UX writing. This conceptual model employs the logic of subtraction to remove unnecessary words, thereby maintaining effective communication and reducing cognitive load for users within intricate digital environments.

1.3 Describe Relevant Scholarship

This research builds upon a comparative analysis of major theoretical frameworks, notably Noam Chomsky's (1995) Minimalist Program, which emphasizes structural economy and absolute functional efficiency, and Ernest Hemingway's Iceberg Theory, which highlights the power of information omission to create depth in literature. With this theoretical foundation in place, early empirical investigations into microcopy have predominantly focused on its cognitive and behavioral implications within digital environments. Pioneering research by Song and Schwarz (2008) introduced the concept of processing fluency, demonstrating that linguistic simplicity and visual clarity significantly reduce the perceived difficulty of tasks, thereby enhancing brand trust. Expanding on this cognitive premise, industry case studies have quantified the behavioral effectiveness of lexical adjustments. For instance, Stanphill (2017) reported a 17% increase in user interaction simply by changing the microcopy from a descriptive "book a room" to a more action-oriented "check availability", highlighting the significant impact of minor textual alterations. Similarly, Nielsen (1996) affirmed that simplifying interface language optimizes scanning efficiency, establishing a direct correlation between linguistic clarity and improved task completion rates. Beyond mere cognitive facilitation, microcopy serves as a powerful socio-pragmatic tool. Portmann (2022) critically examines the "symbolic violence" of these seemingly innocuous texts, asserting that microcopy, such as cookie consent notices, is not just functional but actively shapes user choices and reinforces the systemic power dynamics inherent in digital platforms.

In the realm of linguistic optimization, professional discourse has increasingly drawn structural parallels between UX writing and Ernest Hemingway's literary minimalism. Motahari (2020) highlights parataxis, the juxtaposition of short, independent clauses without conjunctions, as a shared technique that lessens mental effort and echoes the precision characteristic of Hemingway's prose. Supporting this viewpoint, Wong (2020) advocates applying Hemingway's Plain English principles, underscoring that the systematic removal of adverbs and complex syntactic structures is crucial for reducing cognitive load in digital interfaces.

However, despite these valuable insights, a notable theoretical and structural gap persists in the existing literature. Extensive academic studies have traditionally explored Hemingway's Iceberg Theory as a purely literary and aesthetic concept, while research on UX writing has predominantly focused on usability metrics and cognitive ergonomics. Furthermore, discussions that connect Hemingway's stylistics to UX design have mostly been limited to industry blogs and professional commentary like Motahari (2020) and Wong (2020), rather than being grounded in rigorous, cross-disciplinary academic research.

1.4 State Hypotheses and Their Correspondence to Research Design

There is a notable lack of empirical studies that systematically synthesize these two areas. The current study aims to bridge this theoretical gap by formally operationalizing Hemingway's minimalist stylistics into a functional framework for digital product design. By conceptualizing the transition from the artistic Iceberg Theory to an applied Iceberg Standard, this research advocates for a rigorous subtractive methodology in UX writing. Ultimately, this framework responds to the critical demands of the contemporary attention economy, offering a systematic approach to crafting succinct, high-impact microcopy that optimizes cognitive processing while preserving communicative depth.

Building on this foundation, the study proposes the hypothesis that UX writing fosters an implicit information model that fundamentally parallels the principles of the Iceberg Theory. However, the hypothesis also indicates a divergence: while omissions in literature carry emotional significance, the information gaps present in UX writing are strategically designed to minimize cognitive load and facilitate a coherent flow of information for users. To test this hypothesis, the research employs a mixed-methods comparative design. Quantitatively, the study utilizes corpus linguistics through AntConc software to systematically compare a Literary Corpus (comprising Hemingway's major works) with a UX Writing Corpus (derived from Spotify, Slack, Duolingo, and Google). Subsequently, the research integrates these findings with qualitative functional analysis to elucidate the application of the logic of subtraction, demonstrating the transformation of minimalism from a mere aesthetic literary device into an applied digital standard.

2. Method

This research utilizes a mixed-methods comparative design, combining quantitative corpus linguistics with qualitative functional analysis in a systematic manner. The main aim is to investigate the historical functional transition of minimalist language, from its artistic use in 20th-century literature to its practical application in modern digital interfaces.

2.1 Corpus Compilation

To facilitate a comprehensive comparative evaluation, two specifically curated textual datasets have been constructed.

(1) The Literary Corpus (see Appendix 1): This dataset comprises 680 tokens and features carefully selected excerpts from three seminal works by the renowned author Ernest Hemingway. The texts chosen for inclusion are *For Whom the Bell Tolls*, *The Old Man and the Sea*, and *The Sun Also Rises*. These excerpts are meticulously selected to demonstrate Hemingway's distinctive application of the Iceberg Theory, which emphasizes the power of subtext and implicit meaning beneath surface dialogue. A particular focus is placed on localized paratactic dialogue, characterized by succinct, understated exchanges such as Yes or I do not remember it. Additionally, the dataset highlights Hemingway's hallmark unembellished descriptive prose; for instance, phrases like "He lay flat on the brown" (Hemingway, 1940, p.3) and "pine-needed floor are showcased to illustrate his minimalist yet evocative style.

(2) The UX Writing Corpus (see Appendix 2): This dataset is a collection of carefully curated user interface elements sourced from several prominent digital platforms, including Spotify (2026), Slack (2026), Duolingo (2026), and Google AI Studio (2026). The aim of this compilation is to capture functional microcopy that plays a critical role in enhancing user experience. The dataset includes a variety of instructional prompts, such as Start with a simple prompt from Google AI Studio and Tap the pairs from Duolingo, which guide users through specific tasks. Furthermore, it encompasses system feedback mechanisms designed to reassure users, illustrated by phrases like Got it! from Slack. This meticulous gathering of microcopy reflects the nuances of user interactions within digital environments, providing insights into effective communication strategies in UX writing.

2.2 Quantitative Corpus Analysis

The quantitative phase of the study leveraged AntConc, a powerful and sophisticated corpus analysis tool, to conduct a thorough examination and comparison of the lexical structures present within the two datasets being analyzed. The computational evaluation was structured around three crucial metrics that provided insights into the linguistic features of each corpus.

(1) Lexical Diversity (TTR): The TTR was meticulously calculated to evaluate the variety and breadth of vocabulary in relation to the overall length of the texts. This metric serves as an objective measure of the density of unique semantic units within the corpora. The analysis was designed to investigate the hypothesis that UX writing, which often targets functionality and clarity, requires a lexicon that is highly specific and practical. This was contrasted with the rhythmic and sometimes repetitive lexical choices that characterize Ernest Hemingway's

prose, known for its simplicity but depth, thus highlighting the differences in vocabulary usage and stylistic approaches between the two forms of writing.

(2) Frequency Profiling and Pronominal Distribution: To facilitate a deeper understanding of the text's syntactic distribution, frequency word lists were generated for both datasets. This process involved isolating and analyzing grammatical function words, such as prepositions and conjunctions, alongside thematic nouns. For instance, the analysis compared existential thematic nouns typically found in literature, like sea and dead, against functional terminology prevalent in digital contexts, such as data, music, and ad. Moreover, the study tracked shifts in pronominal perspective, which included a focused measurement of the frequencies of the third-person narrative voice (e.g., He/His) in contrast to the second-person, user-centric direct address (e.g., You/Your). This distinction is significant as it reflects differences in the engagement of the audience between literary and UX writing.

(3) Morphological Paradigm Analysis: Utilizing the sorting algorithm TypeEnd feature of AntConc, the investigation delved into verb inflections to decode the primary functional orientation of each corpus. This detailed analysis aimed to contrast narrative past-tense markers, indicative of physical and chronological actions within literature, exemplified by verbs like washed and fished, against morphological constructs that convey specific states, requirements, or updates within digital systems, such as downloaded, needed, and updated. This examination was instrumental in elucidating the distinct functional roles that verb forms play within the different genres, thereby shedding light on how language functions to create meaning in varying contexts.

2.3 Qualitative Functional Analysis

After the quantitative data extraction was completed, a thorough qualitative functional analysis was initiated. The main objective was to delve deeper into the underlying principles of subtraction within both literary and digital contexts. This analytical phase was grounded in three distinct theoretical frameworks, each offering a unique lens through which to examine the role of omission and information structuring.

(1) The Pragmatics of Omission: Drawing inspiration from Hemingway's Iceberg Theory, this study scrutinized the contextual implications of leaving out certain structural elements. The analysis paid particular attention to how the purposeful exclusion of explicit emotional adjectives in literature contributes to a rich subtext that invites diverse interpretations from readers. By creating a narrative space filled with implicit meaning, Hemingway's technique cultivates an environment where the reader engages more deeply with the text. This literary approach was contrasted with the use of microcopy in UX design, where the intentional removal of superfluous instructional language serves to alleviate cognitive burden. In UX writing, this streamlining fosters quicker task completion, allowing users to focus on essential actions without distraction.

(2) Information Flow Architecture: This framework was based on the principles of the Given-New information structure, as thoroughly explained. The analysis here centered on how the sequencing of information influences understanding and retention. By examining Hemingway's method of spatial narrative sequencing, it was noted that he often introduces a character before illuminating the surrounding environment, such as the wind or mountainside. This technique enhances reader immersion through contextual buildup. In contrast, the structured cognitive pacing seen in digital interfaces was highlighted, where platforms like Slack exemplify this concept. For instance, Slack's design isolates the email input stage, prompting users to first establish their workspace name. This deliberate sequencing mitigates the likelihood of overwhelming users with too much information at once, showcasing an effective application of progressive disclosure.

(3) Syntactic Positioning and Modality: Lastly, a comprehensive comparative analysis was conducted to investigate the frequency and implications of imperative sentence structures within UX writing versus literary narrative styles. This segment of the study scrutinized the reliance on direct, action-oriented commands, an approach commonly employed in user interfaces. A prime example is Spotify's assertive prompt, "Search for artists, songs, or podcasts," which directs users with clear and immediate instructions. This stylistic choice contrasts sharply with Hemingway's use of a more declarative and observer-centered narrative voice, which tends to elicit reflection rather than action. This exploration of syntactic positioning reveals how different grammatical choices can significantly influence user engagement and experience in varied contexts.

Through the integration of these three frameworks, the study provided a nuanced understanding of how omission and structural choices impact reader and user interactions, ultimately highlighting the critical role of clarity and context in both literature and digital communication.

2.4 Comparative Synthesis

In the concluding methodological phase of this study, a comprehensive comparative triangulation framework was utilized to integrate and synthesize a range of quantitative metrics and qualitative structural findings. The quantitative metrics examined included various measures such as TTR, detailed frequency lists of linguistic elements, and an analysis of morphological markers. These metrics were carefully selected for their ability to provide objective insights into language usage and patterns.

On the qualitative side, the analysis focused on structural findings that encompass critical elements such as information flow and modality. By examining how information is structured and conveyed within the text, we sought to reveal deeper layers of meaning and intent.

This dual approach is applied with the aim of reinforcing Chomsky's principle of linguistic economy, which posits that language is designed to communicate complex ideas efficiently. By bridging these two domains, quantitative metrics and qualitative analysis, it provides empirical evidence that illustrates how a common mechanism of linguistic subtraction can serve divergent purposes. Specifically, in the realm of literature, this mechanism enhances subtextual meaning and emotional resonance, allowing for a richer reader experience. Conversely, in the context of technology, it focuses on maximizing operational efficiency and cognitive processing abilities, ensuring that information is delivered in the most effective manner possible. This multifaceted exploration ultimately sheds light on the intricate relationship between language, cognition, and the varied functions it serves across different contexts.

3. Results

The analytical stage yielded both quantitative lexical measurements and qualitative structural insights, revealing distinct stylistic differences between literary prose and digital microcopy. Through a comprehensive examination of word frequency, sentence length, and lexical diversity, the authors were able to derive numerical data that quantify the use of language in each format. Additionally, we explored the underlying structures and stylistic choices unique to literary prose, such as metaphor, narrative flow, and character development, contrasting these with the concise, actionable language often found in digital microcopy, which prioritizes brevity and clarity for effective user engagement. This dual approach not only highlighted the varying complexities of each writing style but also underscored the different communicative goals they aim to achieve.

3.1 Quantitative Lexical Analysis

The analysis conducted using AntConc on two distinct datasets, which combined to a total of 1,072 tokens, revealed several noteworthy differences in both morphology and lexicon across the two corpora.

(i) Lexical Diversity (TTR): The UX Writing Corpus, which consists of 388 tokens and 225 unique lexical types, displayed a remarkably high TTR of 57.99%. This elevated TTR indicates a substantial level of lexical diversity within this corpus, suggesting that UX writing employs a rich variety of vocabulary aimed predominantly at facilitating functional navigation. The distinctive terms often encountered include technical and contextual keywords such as API, playlist, and Wi-Fi, which serve specific purposes in guiding users through digital interfaces. Conversely, the Literary Corpus, encompassing 680 tokens and 319 unique types, yielded a notably lower TTR of 46.91%. This reduced diversity implies a different linguistic approach, as seen in the works of Hemingway, where the rhythmic and intentional repetition of simpler, fundamental vocabulary like *old*, *man*, and *sea* is employed. This method serves to cultivate a deeper atmosphere and emotional resonance, highlighting a stylistic choice that prioritizes thematic continuity over varied lexicon.

(ii) Pronominal Orientation: The frequency profiling conducted as part of the analysis revealed a significant divergence in narrative perspectives between the two corpora. The Literary Corpus is primarily characterized by the usage of third-person pronouns, with *He* and *His* ranking as the 7th and 12th most common words, respectively.

This suggests an observational narrative stance typical of literary texts, where the narrator maintains a certain distance from the characters. In stark contrast, the UX Writing Corpus predominantly employs second-person pronouns, with You (ranking 9th) and Your (ranking 13th) taking precedence. This choice reflects a direct, user-centered communicative approach, emphasizing engagement and interaction with the user, which is crucial for effective digital communication and interface design.

(iii) Morphological and Thematic Paradigms: The analysis further utilized the “TypeEnd” sorting algorithm to underscore the distinct functionalities attributed to verbs within each corpus. In the context of literary works, verbs often bear the “-ed” suffix, a morphological marker that typifies chronological or physical past actions. Examples include watched, washed, and fished. This grammatical structure contributes to a narrative flow that anchors the reader in time and event. Conversely, in the realm of UX microcopy, similar morphological forms are utilized to indicate operational states or necessary conditions for digital systems. Terms such as downloaded, updated, and needed highlight a focus on immediacy and functionality, essential for user interaction. The thematic exploration within these corpora also differs markedly; Hemingway’s literary nouns frequently engage with existential concepts, such as sea and dead, imbuing his prose with philosophical depth. In contrast, the nouns found within UX writing tend to be utilitarian and practical, focusing on elements essential for user experience, such as data, music, and ad, reflecting a clear intent to serve functional needs.

Overall, the analysis elucidates how distinct communicative objectives shape the morphological choices and lexical selections in both UX writing and literary endeavors.

3.2 Qualitative Functional Analysis

In addition to analyzing lexical frequencies, the qualitative structural analysis unveiled distinct applications of the logic of subtraction across two different domains: literature and UX writing. This exploration reveals how omission serves varied functions, each tailored to its context.

(i) The Pragmatics of Omission: Ernest Hemingway is renowned for his minimalist style, and his adept use of information gaps exemplifies how omission can shift interpretative responsibility onto the reader. For instance, after the death of a character, Hemingway writes, “*He was dead, and that was all... He felt numb all through himself*” (Hemingway, 1940, p. 442). In this example, the absence of emotionally charged adjectives, such as tragic or painfully, invites readers to engage deeply with the text, encouraging them to confront and analyze the underlying trauma without being explicitly guided through emotional cues. In stark contrast, UX writing employs omission as a tool to minimize cognitive load for users. Take the example from Slack, which streamlines instructions by condensing the phrase “Click here to start sending your message” into a simple “Send” or an easily recognizable arrow icon. This strategy effectively reduces syntactic friction, allowing users to execute tasks more intuitively and swiftly, enhancing the overall usability of the interface.

(ii) Information Flow and Cognitive Pacing: Hemingway’s narrative technique often utilizes a spatial “Given-New” information flow, a structure that grounds the reader in the character’s immediate physical environment before expanding to encompass broader contextual details. For instance, he might first detail a character lying on the floor before revealing the atmospheric elements, such as the wind rustling above. This method effectively immerses the reader in the character’s experience. Conversely, digital interfaces prioritize cognitive pacing to avoid overwhelming users with excessive information. An illustrative case can be found in Slack, which strategically isolates the email input process before introducing secondary tasks related to workspace naming. Similarly, Spotify enhances user engagement by prioritizing personalized expectation-setting features, like “Made for Linh”, rather than delivering generic commands. This personalized touch emotionally guides users through the interface, making their experience more relatable and tailored.

(iii) Syntactic Modality (Imperatives): The analysis highlighted that UX writing frequently relies on imperative structures for clarity and actionability. Unlike traditional command statements that may feel abrupt or demanding, the imperatives employed in UX writing are crafted to serve as gentle but direct guidance. For example, when Spotify prompts users with “Search for artists, songs, or podcasts,” it does so in a way that fosters a sense of

empathy and directness. This approach intentionally omits polite hedging phrases like “You can...” or “Please...,” thus streamlining communication and facilitating a seamless connection between the user and the digital environment. The result is an interaction that feels more personal and intuitive, aligning with users’ expectations for a fluid human-computer relationship.

In summary, both Hemingway’s literary techniques and UX writing exemplify the strategic use of omission, with each domain adapting this approach to serve its unique goals, encouraging deeper engagement in literature and enhancing usability in digital interfaces.

4. Discussion

The synthesized results offer a comprehensive examination of the developmental trajectory of minimalist language, tracing its evolution from a purely aesthetic choice in the 20th century to a vital, functional standard in the fast-paced digital landscape of the 21st century. Initially, minimalist language was valued primarily for its stylistic simplicity and elegance, often exemplified in the works of writers like Ernest Hemingway, who excelled at conveying profound meaning through sparse prose. However, with the advent of the digital age, the role of minimalist language has transformed significantly. In contemporary contexts, particularly within UX design, minimalist language serves a more utilitarian function. Modern UX microcopy, that is, the concise bits of text that guide users through digital interfaces, demands not only brevity but also clarity and effectiveness. This shift highlights a crucial divergence between Hemingway’s literary techniques and current UX practices. While both emphasize conciseness, their objectives differ markedly: Hemingway aimed to evoke emotion and provoke thought through understated narrative methods, whereas contemporary minimalist language focuses on enhancing usability and facilitating seamless interaction in an increasingly complex digital environment. This transformation reflects broader changes in communication, underscoring the importance of functional clarity as users navigate vast amounts of information.

This research extends and elaborates on the theoretical concepts posited by Chomsky (1995), particularly his notions of Linguistic Economy and the Principle of Least Effort. Within a literary context, Hemingway’s strategic use of minimalist techniques, evident through his employment of repetition and deliberate structural omissions, culminates in a TTR of 46.91%. This statistic implies that readers are compelled to engage cognitively, filling the “information gaps” left intentionally by Hemingway (1940, 1952, 2006) with their interpretations and insights into the narrative’s deeper meanings. In contrast, modern UX writing showcases a higher TTR of 57.99%, focusing predominantly on actionable language. This creates a linguistic economy that explicitly aims to reduce the cognitive load on users, aligning closely with Song and Schwarz’s (2008) theories concerning cognitive processing. This approach to digital minimalism leverages both high functional diversity and syntactic brevity, ultimately fostering enhanced processing fluency and prompting immediate actions from users.

Furthermore, the findings strongly support Portmann’s (2022) assertion that microcopy substantially shapes user identity and interaction dynamics within digital contexts. A striking shift can be observed in the use of pronouns, transitioning from a third-person narrative perspective (such as He/His) in literary texts to a direct second-person address (You/Your) in user interfaces. This alteration indicates a significant socio-pragmatic evolution in communication. It effectively transforms the audience from passive, reflective readers, who are tasked with deciphering layers of subtext, to active, engaged users whose immediate needs and interactions are directly aligned with the functionalities of the system, guided by the language employed.

This sense of urgency and direct engagement is further underscored by the syntactic structures used in UX writing, which employs imperative forms at a rate five times greater than those found in literary texts. This prevalence highlights the emphasis on operational effectiveness over descriptive contemplation, showcasing how the language is designed to drive user behavior and facilitate quick decision-making, which echoes the industry observations made by Stanphill (2017) and Kaley (2024). Ultimately, this study delineates the evolution of a literary philosophy into the paradigm recognized as the Iceberg Standard for UX design. Much like Hemingway, who expertly crafted linguistic gaps to obscure profound emotional layers and deepen narrative impact, contemporary digital designers apply a methodical logic of subtraction (Motahari, 2020; Wong, 2020). They aim to conceal multifaceted

algorithmic and structural frameworks behind interfaces that prioritize clarity and user focus. By interlinking applied linguistics with interface design, this research posits that in today's attention-driven economy, microcopy emerges as a crucial structural system where linguistic reduction serves not only as a key driver of operational efficiency but also actively fosters user engagement and involvement.

The findings of this research hold substantial implications across theoretical, practical, and pedagogical domains. Theoretically, this study broadens the scope of comparative stylistics and applied linguistics by illustrating that classical literary frameworks can be effectively adapted to analyze human-computer interaction (HCI). It elevates microcopy from a mere technical artifact to a significant subject for pragmatic and discourse analysis, establishing that the digital interface is a valid arena for academic linguistic inquiry.

From a practical standpoint, the introduction of the "Iceberg Standard" offers UX designers and technical writers a concrete, linguistically grounded methodology. Instead of arbitrarily reducing word counts, practitioners can strategically utilize information gaps, morphological shifts, and imperative structures to decrease cognitive load while purposefully guiding user behavior. Moreover, there are noteworthy pedagogical implications for modern language education and linguistics training. As reading habits and communication increasingly transition toward digital platforms, incorporating the analysis of digital microcopy into curricula can foster greater lexical sensitivity and pragmatic awareness among learners. By examining how meaning is constructed or intentionally omitted in digital contexts, educators can bridge the divide between classical literary stylistics and contemporary functional language use, equipping language learners with a more flexible, real-world understanding of linguistic economy.

Author Contributions: All authors equally contributed to this study: Conceptualization, H.N.L. and T.T.D.; Methodology, H.N.L. and T.T.D.; Software, H.N.L.; Validation, H.N.L. and T.T.D.; Formal Analysis, H.N.L.; Investigation, H.N.L.; Resources, H.N.L. and T.T.D.; Data Curation, H.N.L.; Writing – Original Draft Preparation, H.N.L.; Writing – Review & Editing, H.N.L. and T.T.D.; Visualization, H.N.L.; Supervision, T.T.D.; Project Administration, T.T.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding; the study was conducted independently.

Conflicts of Interest: The authors declare no conflict of interest

Informed Consent Statement/Ethics approval: Not applicable.

Data Availability Statement: The original datasets (comprising the Literary Corpus and the UX Writing Corpus) presented in this study are included in the article and its supplementary material (Appendix). Further inquiries can be directed to the corresponding author.

Acknowledgments: Not applicable. The authors have no additional support to report beyond what is listed in the Author Contributions and Funding sections.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
<https://doi.org/10.7551/mitpress/9780262527347.001.0001>
- Duolingo (2026). *Duolingo: Language & Chess* (Version 7.42) [Mobile app]. App Store. apple.com
- Google (2026). *Google* (Version 328.0) [Mobile app]. App Store. apple.com
- Hemingway, E. (1940). *For whom the bell tolls*. Scribner.
- Hemingway, E. (1952). *The old man and the sea*. Scribner.
- Hemingway, E. (2006). *The sun also rises*. Scribner.

- Kaley, A. (2024, May 8). *UX writing: Study guide*. Nielsen Norman Group. <https://www.nngroup.com/articles/ux-writing-study-guide/>
- Motahari, H. (2020, April 29). *Lessons from Hemingway for UX writers*. Medium. <https://uxdesign.cc/lessons-from-hemingway-for-ux-writers-8d13b4e6d9b0>
- Nielsen, J. (1996, August 1). *International Web Usability*. Nielsen Norman Group. <https://www.nngroup.com/articles/international-web-usability/>
- Portmann, L. (2022). Crafting an audience: UX writing, user stylization, and the symbolic violence of little texts. *Discourse, Context & Media*, 48, Article 100622. <https://doi.org/10.1016/j.dcm.2022.100622>
- Song, H., & Schwarz, N. (2008). If it's hard to read, it's hard to do: Processing fluency affects effort prediction and willingness to engage. *Psychological Science*, 19(10), 986–988.
- Spotify. (2026). *Spotify* (Version 9.1.68) [Mobile app]. App Store. <https://apps.apple.com/us/app/spotify-music-and-podcasts/id324684580>
- Slack (2026). *Slack* (Version 24.08) [Mobile app]. App Store. [apple.com](https://apps.apple.com/us/app/slack/id1051228801)
- Stanphill, M. (2017, May 17). *How Google UX writers help users* [Conference presentation]. Google I/O 2017.
- Wong, J. (2020, January 15). *How to use the Hemingway App to improve your UX writing*. UX Collective. <https://uxdesign.cc/how-to-use-the-hemingway-app-to-improve-your-ux-writing-7d2d3a3c9e6>

Appendices

Appendix A: Literary Corpus

FOR WHOM THE BELL TOLLS

1. "Is that the mill?" he asked.

"Yes."

"I do not remember it."

"It was built since you were here..."

2. "And at the bridge?"

"Always two. One at each end."

"How many?"

"There are more than a hundred..."

3. "He lay flat on the brown, pine-needled floor of the forest, his chin on his folded arms, and high overhead the wind blew in the tops of the pine trees."

4. "The young man... took his glasses from the pocket of his faded, khaki flannel shirt, wiped the lenses with a handkerchief, screwed the eyepieces around..."

5. "He crossed the stream, picked a double handful, washed the muddy roots clean in the current and then sat down again beside his pack and ate the clean, cool green leaves..."

6. "He was dead and that was all. He looked very small, dead, Robert Jordan thought."

7. "I will do it," Robert Jordan had said. "I will do it all right."

8. "I understand."

9. "He worked his arm through the other strap and settled the weight of the pack against his back. His shirt was still wet from where the pack had rested."

THE OLD MAN AND THE SEA

1. "He was an old man who fished alone in a skiff in the Gulf Stream and he had gone eighty-four days now without taking a fish."

2. "The old man was thin and gaunt with deep wrinkles in the back of his neck."

3. "Everything about him was old except his eyes and they were the same color as the sea and were cheerful and undefeated."

4. "He was shivering with the morning cold."

5. "Then he smiled."

6. "The sail was patched with flour sacks and, furled, it looked like the flag of permanent defeat."

7. "He fitted the rope lashings of the oars onto the thole pins and, leaning forward against the thrust of the blades in the water, he began to row out of the harbour in the dark."

8. "The sun rose thinly from the sea and the old man could see the other boats, low on the water and well in toward the shore, spread out across the current."

9. " 'Santiago,' the boy said to him... 'I could go with you again. We've made some money.' "

10. " 'No,' the old man said. 'You're with a lucky boat. Stay with them.' "

11. " 'I remember,' the old man said. 'I know you did not leave me because you doubted.' "

12. " 'Sleep well old man.' "

13. "He was too simple to wonder when he had attained humility. But he knew he had attained it and he knew it was not disgraceful and it carried no loss of true pride."

14. "Once there had been a tinted photograph of his wife on the wall but he had taken it down because it made him too lonely to see it and it was on the shelf in the corner under his clean shirt."

15. "He no longer dreamed of storms, nor of women, nor of great occurrences, nor of great fish, nor fights, nor contests of strength, nor of his wife. He only dreamed of places now and of the lions on the beach."

THE SUN ALSO RISES

1. "We ate dinner at the Cercle and then took a taxi to the Café d'Harcourt. We sat on the terrace and drank fine and watched the people."

2. "I went out and around the corner to the post-office. I got my mail and came out."

3. "The taxi went up the hill, passed the lighted square, then went into the dark, and finally came out onto the lighted Place de la Concorde."

4. "I could not sleep and I lay awake and then I began to think about Brett and all the rest of it went away. I was thinking about Brett and my mind stopped jumping around and started to go in sort of circles with a center. It was a hell of a thing to do."
5. "'Oh, Jake,' Brett said, 'we could have had such a damned good time together.'"
6. "'Yes,' I said. 'Isn't it pretty to think so?'"

Appendix B: AI writing corpus

GOOGLE AI STUDIO

1. Describe your idea
2. Create conversational voice apps: Use the Gemini Live API to give your app a voice and make your own conversational experience.
3. Animate images with Veo: bring images to life with Veo 3. Let users upload a product phot and turn it into a dynamic video ad, or animate a character's portrait.
4. Use Google Search data: Connect your app to real-time Google Search results. Build an agent that can discuss current events, cite recent news, or fact-check information.

SPOTIFY

1. Create your first playlist
2. It's easy, we'll help you
3. Let's find some podcasts to follow
4. We'll keep you updated on new episodes
5. Install app
6. Sign up free
7. Sign up to get unlimited songs and podcasts with occasional ads. No credit card needed
8. "Made for you"
9. "Life Sucks"
10. "Deep Focus"
11. "Brain Food".
12. "Soft instrumentals to help you focus"
13. "New music from someone you love"
14. "Don't miss out on [Artist]'s new single"
15. "Dive back in"
16. "No internet connection. Don't worry, your downloaded music is still here."
17. "We couldn't find that, but we found something similar."

DUOLINGO

1. Stay sharp and focused to keep your hearts. You got this!
2. Great work! Let's make this a bit harder
3. Your Score connects course progress to real-life skills
4. Prove you're a legend
5. Time to create a profile
6. *"These reminders don't seem to be working. We'll stop sending them for now."*
7. "Amazing! 5 in a row!"
8. "This one's tough, but you've got this!"
9. "Mistakes are how you learn"
9. "Tap the pairs"
10. "Select the correct translation"

SLACK

1. All your work. All in one place
2. There's so much you and your team can accomplish in Slack. Start the interactive tour and click on the yellow guides to explore.
3. Bring your team together with Slack.
4. *"You're here! The day just got better."*
5. "What a day."

6. "Taking a deep breath... okay, ready."
7. "This is the very beginning of the....."
8. "You're all caught up. Take a moment to enjoy the silence."
9. "Got it!"
10. "I'm on it"
11. "I'll remind you about this in 1 hour."
12. "I'm sorry, I'm still just a bot."
10. "Something's not right. We're looking into it!"
11. "We couldn't connect to Slack. It might be your Wi-Fi, or it might be us."

Active Versus Avoidant Leadership in Information Technology: Contrasting Emotional-Intelligence Profiles of Transformational, Transactional, and Passive-Avoidant Leaders

Eniola O. Olagundoye¹

¹ Texas Southern University. ORCID ID: <https://orcid.org/0009-0003-8145-6076>

Abstract

Purpose - This article reframes the dissertation's significant and non-significant findings as an active-versus-avoidant leadership pattern. **Methodology** – Using data from 185 IT professional leaders, the study measured transformational, transactional, and passive-avoidant leadership with the MLQ 5X and four EI components with the SSEIT. **Findings** - Transformational leadership predicted all four EI components; transactional leadership predicted three of four; overall EI predicted transformational leadership ($R^2 = .343$, $\beta = .585$) and transactional leadership ($R^2 = .222$, $\beta = .471$); and overall, EI did not significantly predict passive-avoidant leadership. Six of seven original null hypotheses were rejected, while the passive-avoidant hypothesis was not. **Originality/value** - The study turns the non-significant passive-avoidant result into a theoretically meaningful boundary condition, suggesting that EI is associated with active leadership engagement rather than leadership indiscriminately.

Keywords: Passive-Avoidant Leadership, Laissez-Faire Leadership, Emotional Intelligence, Transformational Leadership, Transactional Leadership, Information Technology

1. Introduction

Leadership research often privileges positive styles and treats passive or laissez-faire leadership as a residual category. Yet avoidance, delayed intervention, and the failure to provide leadership are behaviors with organizational consequences, particularly during technology change. Information technology environments create frequent demands for timely decisions, coordination, problem resolution, and employee support. A passive leader can therefore shape outcomes through inaction as much as an active leader shapes them through action.

The dissertation dataset offers a useful contrast. Transformational leadership predicted all four emotional-intelligence components. Transactional leadership predicted three of four. Overall, EI predicted both transformational and transactional leadership. In contrast, EI did not significantly predict passive-avoidant leadership. Of the seven original null hypotheses, the first six were rejected and only the seventh - the EI/passive-avoidant relationship - was not rejected.

This article argues that the pattern is theoretically more useful when framed as a distinction between active and avoidant leadership. The purpose is not to elevate a null result into proof of no relationship in all settings. Rather, the non-significant finding defines a boundary in this sample and helps clarify that EI was selectively associated with leadership forms requiring active engagement.

2. Theoretical Background

2.1 Active leadership

Transformational and transactional leadership differ in their motivational mechanisms, but both require leaders to engage followers. Transformational leaders articulate vision, stimulate new thinking, model values, and develop followers. Transactional leaders clarify expectations, exchange rewards for performance, monitor standards, and intervene when deviations occur. Both styles therefore involve attention, communication, and behavioral response. Emotional intelligence may support these activities by helping leaders perceive the affective context, regulate their own reactions, manage followers' emotions, and use emotional information in problem solving. Contemporary reviews indicate that leader EI is associated with a range of leadership behaviors and effectiveness outcomes (Gerhardt et al., 2026). In digital transformation, EI may also help leaders balance task, relation, change, and external-oriented behaviors (Orkamo et al., 2025).

2.2 Passive-avoidant leadership

Passive-avoidant leadership includes laissez-faire behavior and passive management-by-exception. It is characterized by delayed decisions, limited intervention, avoidance of responsibility, or action only after problems become serious. Although historically conceptualized as an absence of leadership, recent research treats laissez-faire leadership as an active research problem because it can influence employee experience and change readiness. Groulx et al. (2024), for example, found that manager emotional exhaustion was associated with higher laissez-faire leadership, which in turn undermined team psychological safety and readiness for change. Zheng and Li (2024) describe laissez-faire leadership as potentially double-edged: the absence of guidance may be experienced as resource loss but can also be interpreted as autonomy under some conditions. These findings suggest that passive leadership is not psychologically empty; its antecedents and consequences deserve separate analysis.

2.3 Why emotional intelligence may differentiate active from avoidant leadership

EI is behaviorally relevant when a leader attends to followers, interprets emotional cues, communicates expectations, coaches employees, resolves conflict, or regulates personal reactions. These processes are integral to active leadership. Passive-avoidant leadership, by definition, reduces the frequency of such engagement. It is therefore plausible that self-reported EI will be more strongly associated with active than avoidant styles.

This proposition is especially relevant in technology settings. Digital transformation requires leaders to create acceptance, develop digital competence, and support innovation (Orkamo et al., 2025). Digital leadership reviews similarly emphasize vision, motivation, collaboration, and adaptation (Lin, 2025; Keisham & Lathabhavan, 2026). Avoidant leadership may be particularly poorly matched to these demands.

3. Method

The study used a quantitative, correlational design to examine relationships between self-reported leadership styles and emotional intelligence among information technology professional leaders. Participants were recruited through convenience and snowball sampling from the researcher's professional IT network. A total of 230 surveys were collected; 45 incomplete surveys were excluded, resulting in 185 usable cases.

Leadership style was measured with 36 items from the Multifactor Leadership Questionnaire (MLQ 5X), representing transformational, transactional, and passive-avoidant leadership. Emotional intelligence was measured with the Schutte Self-Report Emotional Intelligence Test (SSEIT), organized in the study into four

components: perception of emotion, managing own emotions, managing others' emotions, and utilization of emotion. Demographic controls included gender, age, ethnicity, education, and tenure in leadership roles within information technology.

The original analyses used blockwise multiple regression at a significance level of $p < .05$. Demographic variables were entered before the focal leadership or emotional-intelligence variables so that relationships of primary interest could be evaluated after accounting for the controls. The design is cross-sectional and based on self-report; consequently, the word predictor refers to statistical prediction and should not be interpreted as proof of causality.

4. Results

Across the leadership-to-EI models, transformational leadership significantly predicted perception of emotion, managing own emotions, managing others' emotions, and utilization of emotion. Transactional leadership significantly predicted perception of emotion, managing others' emotions, and utilization of emotion, but did not predict managing own emotions. Passive-avoidant leadership did not emerge as a comparable significant predictor. In the reverse-direction models, overall emotional intelligence significantly predicted transformational leadership ($R^2 = .343$, $\beta = .585$, $p < .05$) and transactional leadership ($R^2 = .222$, $\beta = .471$, $p < .05$). Overall emotional intelligence did not significantly predict passive-avoidant leadership.

At the hypothesis level, six of the seven original null hypotheses were rejected. The only hypothesis that was not rejected was the proposition of no relationship between overall emotional intelligence and passive-avoidant leadership. The convergence of results in both analytic directions supports an active-versus-avoidant interpretation.

5. Discussion

The study's most distinctive contribution is the pattern across results rather than any isolated coefficient. Emotional intelligence was consistently associated with transformational leadership and, to a somewhat narrower extent, transactional leadership. It was not significantly associated with passive-avoidant leadership. This suggests that EI may be tied to the interpersonal and self-regulatory processes involved in active leadership engagement.

The pattern also distinguishes transformational from transactional leadership. Transactional leadership did not predict managing own emotions, while transformational leadership did. Thus, active leadership is not a single psychological category. Transformational leadership appears to have the broadest emotional profile in this sample. The passive-avoidant null result should be interpreted carefully. A non-significant relationship does not prove that EI is irrelevant to all forms of passive leadership. Contemporary research indicates that emotional exhaustion can contribute to laissez-faire behavior (Groulx et al., 2024), suggesting that other emotional variables may matter. The present result instead shows that overall self-reported EI did not statistically predict passive-avoidant leadership in this sample.

This distinction is valuable for digital transformation. Orkamo et al. (2025) show that successful transformation requires a balance of multiple leadership behaviors, while current digital-leadership research emphasizes adaptation, collaboration, talent development, and strategic action (Lin, 2025; Keisham & Lathabhavan, 2026). These demands are fundamentally active. The study's findings suggest that emotional competencies may be part of the human capability supporting such engagement.

6. Implications for Technology Organizations

Organizations should assess leadership readiness among technical professionals not only in terms of expertise but also in terms of behavioral engagement and emotional competence. Leadership development can focus on timely decision making, constructive intervention, communication, coaching, emotion regulation, and conflict management.

Passive-avoidant tendencies deserve explicit attention in leadership assessment. In technology projects, delayed decisions or limited support can contribute to unresolved issues, uncertainty, and loss of confidence. The original study did not measure project performance, cybersecurity outcomes, or implementation success, so those consequences should not be claimed from the data. They are testable implications for future research.

AI adoption raises an additional question: whether leaders become more or less passive when algorithmic systems provide recommendations or automate managerial tasks. Future studies should test whether reliance on AI can unintentionally encourage leadership withdrawal, or whether emotionally intelligent leaders use AI while maintaining human engagement and accountability.

7. Limitations and Future Research

The study is cross-sectional, self-reported, and based on a convenience sample, which limits causal inference and generalizability. The passive-avoidant score was relatively low on average, which may also constrain variance. The study did not distinguish the passive management-by-exception and laissez-faire subdimensions in the focal article-level analysis.

Future research should examine the passive components separately, use follower ratings, and include outcomes such as psychological safety, readiness for change, project performance, technostress (Ertiö et al., 2024), and employee retention. Longitudinal designs could test whether EI development reduces passive-avoidant behaviors or whether emotional exhaustion predicts movement toward leadership withdrawal. AI-enabled management environments offer an especially important context for examining whether technology augments active leadership or facilitates avoidance.

8. Conclusion

The dissertation results reveal a coherent boundary: emotional intelligence was associated with transformational and transactional leadership but not with passive-avoidant leadership. Reframing the findings as active versus avoidant leadership produces a distinct theoretical contribution and gives meaning to the one null hypothesis that was not rejected. The pattern is especially relevant to contemporary technology organizations, where digital transformation and AI require leaders to remain behaviorally and emotionally engaged even as technology assumes a larger role in decision support and work execution.

Data and Publication Note: This article was developed from the author's dissertation dataset and uses the same underlying sample and measures while addressing a distinct research question and theoretical contribution. Prior dissemination of the dissertation and related publications from the same dataset has been disclosed in accordance with the journal's publication policies.

Funding: This study received no funding.

Conflict of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics Approval: The original dissertation study was approved by the Our Lady of the Lake University Institutional Review Board (OLLU IRB) before data collection and was designated EXEMPT STATUS on February 15, 2016. Informed consent was presented to each respondent before the online survey. Participation was voluntary, respondents could discontinue at any time without penalty, and anonymity was maintained because names were not collected or linked to survey results.

References

Bass, B. M. (1985). *Leadership and performance beyond expectations*. Free Press.

- Bass, B. M., & Avolio, B. J. (1994). Improving organizational effectiveness through transformational leadership. Sage.
- Bass, B. M., & Avolio, B. J. (2004). Multifactor Leadership Questionnaire manual and sample set (3rd ed.). Mind Garden.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–31). Basic Books.
- Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9, 185–211.
- Schutte, N. S., Malouff, J. M., & Bhullar, N. (2009). The assessing emotions scale. In C. Stough, D. H. Saklofske, & J. D. A. Parker (Eds.), *Assessing emotional intelligence* (pp. 119–134). Springer.
- Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, 25(2), 167–177.
- Thite, M. (2000). Leadership styles in information technology projects. *International Journal of Project Management*, 18(4), 235–241.
- Wong, C. S., & Law, K. S. (2002). The effects of leader and follower emotional intelligence on performance and attitude: An exploratory study. *The Leadership Quarterly*, 13(3), 243–274.
- Gerhardt, K., Bauwens, R., & van Woerkom, M. (2026). Emotional intelligence and leader outcomes: A comprehensive review and roadmap for future inquiry. *Human Resource Development Review*, 25(1). <https://doi.org/10.1177/15344843251342689>
- Groulx, P., Maisonneuve, F., Harvey, J.-F., & Johnson, K. J. (2024). The ripple effect of strain in times of change: How manager emotional exhaustion affects team psychological safety and readiness to change. *Frontiers in Psychology*, 15, 1298104. <https://doi.org/10.3389/fpsyg.2024.1298104>
- Zheng, S., & Li, L. (2024). Navigating the double-edged sword: How and when laissez-faire leadership affects employees' challenging behavior. *Acta Psychologica*, 250, 104549. <https://doi.org/10.1016/j.actpsy.2024.104549>
- Orkamo, M., Ukko, J., Rantala, T., & Saunila, M. (2025). Leadership behaviours to promote organisational performance in private sector digital transformation: A systematic literature review. *Digital Business*, 5(2), 100155. <https://doi.org/10.1016/j.digbus.2025.100155>
- Lin, Q. (2025). Digital leadership: A systematic literature review and future research agenda. *European Journal of Innovation Management*, 28(6), 2469–2488. <https://doi.org/10.1108/EJIM-07-2023-0522>
- Keisham, B. D., & Lathabhavan, R. (2026). Digital leadership in the age of AI: A review. *Development and Learning in Organizations: An International Journal*, 40(2), 14–17. <https://doi.org/10.1108/DLO-06-2025-0219>
- Ertiö, T., Eriksson, T., Rowan, W., & McCarthy, S. (2024). The role of digital leaders' emotional intelligence in mitigating employee technostress. *Business Horizons*, 67(4), 399–409. <https://doi.org/10.1016/j.bushor.2024.03.004>

Bringing Essentialism in 21st-Century Education: Exploring the Thoughts of William C. Bagley

Marselus Ruben Payong¹

¹ Universitas Katolik Indonesia Santu Paulus Ruteng. ORCID ID: <https://orcid.org/0000-0002-3229-3309>

Abstract

William Chandler Bagley (1874–1946) was an American educational figure known as a pioneer of essentialism. His ideas emphasized the importance of the core curriculum, teacher professionalism, and intellectual discipline in shaping students' character and thinking skills. Although born in the context of the early 20th century, Bagley's ideas remain relevant to education in the 21st-century, characterized by globalization, digitalization, and rapid socio-cultural change. This article aims to analyze the influence and relevance of Bagley's ideas on contemporary educational practice by examining their relationship to curriculum, teacher education, and the 21st-century learning paradigm. The method used is a historical-critical literature review of Bagley's primary works and modern secondary literature. The study results show that Bagley's principles, especially the emphasis on curriculum content, academic standards, and teachers' professional authority—remain an important foundation of modern education, although they need to be reformulated to be compatible with 21st-century values of democracy, creativity, and inclusivity.

Keywords: Essentialism, Core Curriculum, Intellectual Discipline, 21st Century Skills

1. Introduction

Modern education faces increasingly complex challenges: technological acceleration, economic globalization, social transformation, and a crisis in the values and meaning of learning. In this context, debates about the direction and nature of education have resurfaced, particularly regarding the tension between individual freedom, creativity, and the need for measurable academic standards. Amidst the increasingly dominant trend of educational progressivism in the 21st-century paradigm, interest has resurfaced in the classical ideas about the essence of education, intellectual discipline, and teacher professionalism, as articulated by William Chandler Bagley (1874–1946)—a leading figure in the essentialist school of American educational philosophy (Null, 2007; Sahin, 2018). William C. Bagley was a philosopher and educator who opposed the excesses of educational progressivism that developed in the early 20th century, particularly the ideas of John Dewey and his students such as William H. Kilpatrick. In Bagley's view, education should not simply be a social laboratory that prioritizes the freedom of students without direction, but rather an institution that transmits values, intellectual habits, and knowledge proven in human culture (Bagley, 1931). He emphasized that the primary task of schools is to maintain the continuity of civilization by transmitting essential knowledge to the younger generation (Bagley, 1926; Bestor, 1953).

His pioneering idea of essentialism was a reaction to progressivism, which overemphasized individual experience and student free activity. For Bagley, proper education must include elements of discipline, perseverance, and mastery of academic content, because without a strong intellectual foundation, freedom of thought becomes shallow and irresponsible (Cremin, 1961; Null, 2003). This view made him one of the most important figures in the history of American educational philosophy and the initiator of a new direction emphasizing the balance between content, process, and social values in education.

Although Bagley's ideas emerged over a century ago, their relevance to 21st-century educational practice is resurfacing. Amidst the demands of developing 21st-century skills—such as critical thinking, collaboration, creativity, and technological literacy, education systems still require a robust curriculum structure, professional teachers, and an academic culture that values intellectual perseverance (Trilling and Fadel, 2009; OECD, 2019). Thus, an important question arises: to what extent can Bagley's classical thinking provide a new direction for modern education, faced with the challenges of digitalization, globalization, and cultural pluralism?

Furthermore, the current education system faces a dilemma between two extremes: on the one hand, constructivist and progressive tendencies that emphasize flexibility and personalized learning; on the other, the need for accountability, objective evaluation, and a national curriculum that demands uniform standards (Ornstein and Hunkins, 2018). In this context, Bagley's essentialist approach can offer a middle ground—emphasizing the importance of academic content and the integrity of teacher professionalism, without rejecting pedagogical innovations that meet the demands of the times (Sahin, 2018; Dunn, 2005).

From a historical perspective, Bagley brought classical ideals about education and the practical needs of modern industrial society (Gutek, 2011). He believed that education should be oriented toward developing individuals with character, capable of systematic thinking, and ready to contribute to social progress. These values are now being sought again in an era dominated by project-based learning, digital media, and artificial intelligence algorithms that sometimes neglect the ethical dimensions and substance of scientific knowledge (UNESCO, 2021; Zhao, 2012).

Therefore, this study is crucial for reflecting on William C. Bagley's thinking within the modern philosophy of education and examining its relevance for 21st-century educational practices, which demand a combination of core knowledge, creativity, and national character. Through a theoretical-historical examination of Bagley's work and ideas, this article attempts to build a conceptual bridge between classical essentialism and the contemporary transformative learning paradigm.

Based on this background, three main research questions can be formulated: 1) What are the main points of William C. Bagley's educational thinking that form the basis of the essentialism school? 2) How does Bagley's thinking influence educational practice, particularly in curriculum development and teacher education? 3) What is the relevance and potential reinterpretation of essentialism principles in the context of 21st-century learning?

This research aims to: 1) Comprehensively describe the main principles of William C. Bagley's thinking, including his views on the nature of education, the role of teachers, and curriculum structure. 2) Analyze the influence of Bagley's thinking on the development of educational theory and practice from the early 20th century to the contemporary era. 3) Evaluate the relevance of essentialist principles to the 21st-century educational paradigm and propose a conceptual reinterpretation to suit the global, digital, and multicultural context.

This research is significant in three dimensions: 1) Historical dimension: it provides a deeper understanding of William C. Bagley's role as a key figure in shaping the modern educational paradigm and his contribution to the birth of the back-to-basics movement and core curricula in various countries (Bestor, 1956; Hirsch, 1996). 2) Theoretical dimension: it expands the study of the relationship between essentialism and 21st-century learning theories and opens a space for synthesis between classical values and the demands of modern education. 3) Practical dimension: it offers reflections for curriculum designers, educators, and policymakers on how to balance the depth of academic content, pedagogical creativity, and social inclusiveness in the educational process.

2. Research Method

This research uses a qualitative approach with historical-philosophical inquiry design. More specifically, this research combines two analytical frameworks: 1) Historical analysis, to explore the social, political, and intellectual contexts underlying Bagley's thinking. 2) Philosophical-critical analysis, to examine the conceptual structure, ontological, epistemological, and axiological assumptions in Bagley's thinking, and their relevance to modern educational practice (Knight, 2008; Ozmon and Craver, 2003).

This research utilizes secondary data obtained through in-depth library research. The data sources consist of 1) William C. Bagley's primary works, including: *The Educative Process* (1905); *Classroom Management* (1907); *Educational Values* (1922), *Education and Emergent Man* (1934); and *Determinants of Social Education* (1931). These works serve as the primary material for reconstructing Bagley's original views on the nature of education, teachers, and curriculum. Data analysis was conducted using thematic content analysis, which focuses on interpreting philosophical and historical meanings (Bowen, 2009; Braun and Clarke, 2023).

The analysis process involved four main stages: 1) Description – presenting the historical and biographical context of William C. Bagley and a map of essentialist thought. 2) Interpretation – outlining the key meanings of Bagley's ideas on education, teachers, and curriculum. 3) Critique – evaluating the conceptual strengths and weaknesses of Bagley's thinking based on the framework of modern educational theory. 4) Recontextualization – reinterpreting Bagley's ideas within a 21st-century framework, including their integration with the concepts of 21st-century skills, digital literacy, and global education.

3. Results and Discussion

3.1 Intellectual Biography of William C. Bagley

William Chandler Bagley (1874–1946) was a central figure in American educational philosophy in the early 20th century. He earned his undergraduate degree from the University of Michigan and then pursued graduate studies at Cornell University, where he developed a deep interest in educational psychology and the philosophy of learning (Null, 2003). Having taught at several major universities, including Montana State College and the University of Illinois, Bagley became known as the founder and advocate of essentialism—a reaction to the dominance of progressivism in education (Cremin, 1961; Gutek, 2011).

His academic career flourished during a time when educational progressivism, particularly the ideas of John Dewey and William H. Kilpatrick, dominated American pedagogical discourse. However, Bagley recognized the dangers of overemphasizing “learning through experience” without a strong academic discipline. He wrote several monumental works, such as *The Educative Process* (1905), *Education and Emergent Man* (1934), and *Determinants of Social Education* (1931), all of which emphasized the importance of curriculum content, mental discipline, and teacher professionalism as the core of the educational process (Bagley, 1934; Null, 2007).

In a socio-cultural context, Bagley lived during a time of rapid American industrialization. The workforce required workers who were not only technically skilled but also possessed systematic thinking skills and a strong work ethic. This led to his idea that education should be a means of systematic intellectual and moral training. He rejected the view that education should simply follow a child's interests, as for Bagley, the task of education was to guide, not merely facilitate (Bagley, 1931; Null, 2003).

3.2 Core Principles of Essentialism

3.2.1 Education as the Transmission of Values and Knowledge

Bagley emphasized that education's primary function is as a vehicle for cultural transmission and the transmission of fundamental values. He rejected the progressivist view, which overemphasized social change and

experimentation without considering the continuity of values (Bagley, 1934; Bestor, 1953). He argued that humans become civilized because they inherit the accumulated results of past experiences and proven knowledge. Without this, each generation would start from scratch.

Within this framework, the curriculum must contain "core knowledge" that shapes students' intellectual structure, not merely temporary practical skills (Hirsch, 1996; Ornstein and Hunkins, 2018). This view later became the basis for the core curriculum and back-to-basics movements in the United States in the mid-20th century, particularly in response to concerns about declining academic standards (Bestor, 1956; Ravitch, 2016).

3.2.2 Intellectual Discipline and Perseverance

In contrast to the progressive view that viewed learning as a natural and enjoyable activity, Bagley emphasized intellectual discipline and perseverance as prerequisites for developing mature character and thinking skills (Bagley, 1905). He argued that true learning is not always enjoyable; it is in the face of intellectual difficulties that humans learn to think systematically and responsibly (Null, 2007).

This concept rejects the idea of "learning without coercion" popular among progressivists. For Bagley, teachers must be willing to demand effort and perseverance from students, because without this, character development and critical thinking skills will not be achieved (Cremin, 1961; Gutek, 2011).

3.2.3 Teacher Professionalism

One of Bagley's major contributions was his assertion that teachers are intellectual professionals, not simply learning facilitators. In his work, *Classroom Management* (1907), he emphasized the need for scientific and moral training for prospective teachers to enable them to manage classrooms, transmit knowledge, and shape students' characters (Bagley, 1907; Null, 2003).

Bagley later founded the *Society for the Advancement of Education* as a platform to counter extreme progressivism and to champion scientifically based teacher professionalism. In the 21st-century context, this idea remains relevant as the teaching profession faces new challenges due to the digitalization and automation of the learning process (OECD, 2019; Darling-Hammond, 2020).

3.2.4 Education and Citizenship

Bagley believed that education not only prepares individuals for work but also shapes citizens with a sense of social responsibility. In *Education and Emergent Man* (1934), he wrote that the goal of education is to prepare cultured individuals capable of maintaining a balance between personal and social interests. This principle emphasizes the moral and social dimensions of education, which are now re-elevated in the concepts of education for sustainable development (UNESCO, 2021) and global citizenship education (Banks, 2015).

3.3 Essentialism vs. Progressivism

The fundamental difference between Bagley and progressives like Dewey or Kilpatrick lies in their views on the nature of learning and the role of teachers. For Dewey, learning is the social reconstruction of experience; for Bagley, learning is the systematic internalization of knowledge that has been tested by civilization (Dewey, 1916; Bagley, 1931).

Progressives argue that students learn best through hands-on activities relevant to their lives (learning by doing). However, Bagley believes this view ignores the reality that not all knowledge can be acquired through direct experience; much conceptual knowledge requires explicit instruction, practice, and disciplined thinking (Null, 2003; Sahin, 2018).

On the other hand, Bagley does not completely reject progressivism. He values the role of motivation and relevance in learning but emphasizes the need for a stable and hierarchical curriculum structure so that students have a solid intellectual foundation before exploring independent thinking (Ornstein & Levine, 2008). Thus, Bagley's position can be understood as moderate essentialism—a synthesis between the need for academic content and individual development.

3.4 The Influence of Bagley's Thinking on Modern Education

3.4.1 The "Back to Basics" Movement and the Core Curriculum

Bagley's thinking formed the basis for the "Back to Basics" movement that emerged in America in the 1950s–1970s, in response to concerns about the decline in educational quality and the lack of mastery of basic knowledge (Bestor, 1956). This movement emphasized the importance of literacy, numeracy, history, and science as the main pillars of general education.

The concept of the core curriculum, developed by E. D. Hirsch (1996) and reinforced by research on cultural literacy (Hirsch (1987), is also directly rooted in Bagley's idea that there is universal knowledge that every citizen must possess to participate in a democratic society.

3.4.2 Teacher Education and Academic Professionalism

Bagley also left an important legacy in the formation of teacher education institutions. He believed that teachers should be intellectuals trained in both academic and pedagogical fields (Null, 2007). His ideas inspired the emergence of university-based teacher preparation programs that emphasize a balance between theory and practice.

In the 21st century, this idea has taken on new forms through the professional learning communities (PLC) model, reflective teaching, and teacher inquiry, all of which seek to affirm the role of teachers as professional learners (Hord, 2004; Darling-Hammond, 2021).

3.4.3 Relevance to 21st-Century Education

In the digital era, Bagley's essentialism seems to take on new meaning. Basic knowledge now consists not only of memorizing facts but also of information, data, and media literacy (Trilling & Fadel, 2009). The principle of core knowledge translates into 21st-century literacies, which demand a deep understanding of science, technology, culture, and digital ethics (OECD, 2019; Zhao, 2012).

Furthermore, the role of teachers as guardians of academic standards has become increasingly important amidst the flood of unverified information. In this context, Bagley's ideals about the intellectual discipline and moral responsibility of teachers have regained their relevance (Sahin, 2018; Schouten & Brighthouse, 2015).

3.5 Criticism and Reinterpretation of Essentialism

Despite its significant influence, Bagley's views have not escaped criticism. Critical pedagogy theorists such as Paulo Freire (1970) believe that the essentialist approach has the potential to perpetuate social power structures by making the teacher the sole authority. Furthermore, approaches that overemphasize academic content are considered to neglect the social context and life experiences of students (Giroux, 2025).

However, modern reinterpretations of essentialism demonstrate that Bagley's principles can be contextualized without losing their essence. For example, the knowledge-rich curriculum model in the UK (Young, McNamara, and Coughlan, 2017) combines the concept of powerful knowledge with an inclusive, reflective approach. Similarly, in the context of digital learning, intellectual discipline can be developed through scaffolded learning and guided inquiry that still respect the structure of knowledge (Kirschner et al., 2006; Fullan and Quinn, 2020).

Thus, Bagley's essentialism retains conceptual viability when reinterpreted within an educational framework that balances knowledge, skills, and character.

3.6. Reactualization of Bagley's Ideas in the Context of Contemporary Education

William C. Bagley's thinking, particularly through the concept of Essentialism, has experienced a revitalization in the 21st century with the increasing need for competency standards, learning effectiveness, and accountability in global education systems (Ornstein & Hunkins, 2018). In an education world marked by the development of digital technology and the student-centered learning paradigm, many thinkers believe that Bagley offers a balance between student freedom and the importance of intellectual discipline (Knight, 2008).

The reactualization of Bagley's ideas is evident in education policies in various countries, particularly the United States and East Asia, which emphasize the core curriculum and outcome-based evaluation. This approach aligns with the essence of essentialism, which places mastery of the fundamentals of knowledge as the foundation for individual and national progress (Bagley, 1934; Ravitch, 2013). In Indonesia, for example, the Pancasila Student Profile and the Independent Curriculum policies can be interpreted as efforts to balance learning freedom with mastery of basic competencies, reflecting the dialectic between progressivism and essentialism (Hamzah et al., 2022; Savitri et al., 2022).

3.6.1. The Role of Teachers in Bagley's Essentialism Perspective

One of Bagley's most significant contributions to educational practice was his emphasis on the role of teachers as intellectual and moral authorities (Bagley, 1938; Brameld, 1956). In the 21st-century context, this role has been reinterpreted: teachers remain central to the formation of students' academic character, but with new capabilities in integrating technology, global collaboration, and 21st-century competencies (Trilling & Fadel, 2009; Voogt & Roblin, 2012).

Recent research shows that teachers who combine an essentialist approach—emphasizing cognitive discipline, basic literacy, and moral responsibility—with an interactive digital approach are more successful in improving students' critical literacy (Fullan and Quinn, 2020; OECD, 2026). This demonstrates the relevance of Bagley's principle that "true education demands directed and sustained mental exercise" (Bagley, 1934, p. 28).

3.6.2. Core Curriculum and Academic Standards

Bagley's core curriculum emphasizes that every student must master a set of general knowledge that forms the basis for thinking and participating in social life (Bagley, 1934; Ravitch, 2016). This idea has regained prominence in the 21st century through various standard-based education reforms and the back-to-basics movement.

Curricula that emphasize basic literacy, numeracy, and science competencies are now a priority for UNESCO and the OECD in addressing global challenges such as disinformation and educational inequality (UNESCO, 2021; OECD, 2026). Although education is increasingly oriented toward 21st-century skills such as creativity and collaboration, Bagley's thinking continues to emphasize that critical thinking skills cannot develop without a strong foundation of knowledge.

Furthermore, learning outcomes-based assessment systems and national assessments in various countries reflect Bagley's understanding of the importance of objectively measuring learning outcomes as evidence of educational effectiveness (Kliebard, 1992). Criticism of this policy, which is considered to overemphasize cognitive aspects and neglect affective aspects—does not negate Bagley's epistemological value: that education must have a measured and rational intellectual direction (Bagley, 1939; Ravitch, 2013).

3.6.3. Intellectual Discipline and Character Building

Bagley believed that education is not merely a tool for adapting to the social environment, but also a means of developing character and moral virtues (Bagley, 1905; Brameld, 1956). In 21st-century education, which

emphasizes learner autonomy, this idea has been challenged but also finds new support from the perspectives of character education and social-emotional learning (Elias, 2019; Lickona, 2004).

Global research shows that education that combines intellectual discipline with character building produces students who are more resilient, possess social empathy, and can adapt to change (Duckworth and Yeager, 2015; Taylor et al., 2017). This approach is in line with Bagley's vision of "moral intelligence growing out of intellectual diligence" (Bagley, 1934, p. 112).

3.7. Critique of Essentialism and Challenges

Despite Bagley's renewed relevance, the essentialist approach has faced considerable criticism, particularly from constructivists and progressivists (Dewey, 1938; Freire, 1970; Giroux, 2011). The primary criticisms focus on the view that essentialism has the potential to stifle creativity, ignore social context, and reinforce hierarchies in education (Apple, 2011; Kincheloe, 2008).

However, modern reinterpretations of Bagley position him not as the antithesis of progressivism, but as a corrective to the excesses of directionless freedom in education (Ornstein and Hunkins, 2018). Thus, a 21st-century approach demands an integration of essentialism and constructivism—producing a synthesis that balances mastery of core concepts with the ability to think creatively (Darling-Hammond et al., 2019).

The challenge of implementing Bagley's thinking in the digital age lies in how educators can maintain depth in a fast-paced and shallow learning environment (Postman, 1985; Zhao, 2012). Therefore, 21st-century educational reforms that wish to remain grounded in fundamental values must combine Bagley's essentialist tradition with contemporary pedagogical innovations such as blended learning, inquiry-based learning, and AI-assisted instruction (Selwyn, 2022; Luckin, 2023).

3.8. Synthesis: The Enduring Relevance of Bagley's Thinking

This discussion demonstrates that William C. Bagley's ideas remain relevant in the global effort to create an education that is deeply rooted yet open to change. Essentialism in its modern form is no longer simply a conservative movement, but rather an epistemological strategy to ensure that education does not lose its intellectual and moral orientation amidst the rapid flow of technology and value relativism (Bates, 2024).

Emphasizing the balance between core knowledge, intellectual discipline, and character development, Bagley's thinking offers a solid philosophical framework for transforming 21st-century education. He reminds us that true progress is possible only when freedom to learn is guided by meaningful knowledge structures and clear moral goals.

4. Conclusions and Recommendations

4.1. Conclusion

William C. Bagley's thinking has philosophical and practical significance that persists into the 21st century. Through his essentialism, Bagley emphasized the importance of education based on mastery of basic knowledge, intellectual discipline, and the formation of moral character. In Bagley's view, true education not only prepares individuals to adapt to social change but also develops human rational and ethical capacities to responsibly navigate that change (Bagley, 1934).

The analysis shows that Bagley's influence is still felt in various modern educational policies and practices, such as the emphasis on the core curriculum, standards-based evaluation, and the teacher quality improvement movement. The principle of essentialism has proven to be an epistemological pillar for educational systems that strive to maintain a balance between pedagogical innovation and the stability of fundamental values.

In the 21st-century context marked by digitalization, globalization, and value fragmentation, Bagley's thinking serves as a reminder of the importance of disciplined thinking and moral cohesion in the educational process. His idea of the teacher's role as guardian of intellectual and moral quality remains highly relevant in shaping a generation of learners who are not only technologically savvy but also possess a depth of knowledge and social responsibility (Fullan & Quinn, 2020; Lickona, 2004; OECD, 2026).

Thus, William C. Bagley's thinking remains the foundation for developing a balanced educational paradigm: between learning freedom and curriculum structure, between creativity and discipline, and between adaptation and moral principles. Revitalized essentialism is not a form of conservatism, but rather an epistemological strategy to ensure the sustainability of intellectual quality amidst rapid and complex change.

4.2. Recommendations

4.2.1. Reinterpreting Essentialism in 21st-Century Education

Educational institutions need to reinterpret essentialism within a 21st-century educational framework oriented toward global competencies and digital ethics. This means that mastery of the "core" (literacy, numeracy, and basic science) must be integrated with critical, creative, collaborative, and communicative thinking skills (the 4C skills) (Trilling & Fadel, 2009).

4.2.2. Strengthening Teacher Professionalism

Teachers in the 21st century need to take on a dual role as facilitators and guardians of academic quality, as emphasized by Bagley. Therefore, teacher training must emphasize not only technological mastery but also the ability to instill intellectual and moral values (Elias, 2019; Voogt and Roblin, 2012).

4.2.3. Values-Based Curriculum and Core Knowledge

National and local curricula should integrate core knowledge with the social and cultural contexts of students. This aligns with Bagley's spirit of developing an education that is rooted yet adaptive, establishing a balance between tradition and innovation.

4.2.4. Holistic Evaluation

The evaluation system needs to combine cognitive assessment with moral, social, and character indicators. This approach reflects Bagley's principle that education should not stop at mastering content, but rather should shape the whole person (Bagley, 1939; Lickona, 2004).

4.2.5. Collaboration between Tradition and Innovation

Today's education must utilize technology and artificial intelligence without abandoning the core values formulated by Bagley: hard work, disciplined thinking, and social responsibility (Selwyn, 2022; Luckin, 2025). The integration of technology and humanism is the foundation for sustainable and civilized education.

4.3. Implications for Educational Research and Policy

Bagley's thinking opens ample scope for further research in educational philosophy and policy. Some potential research areas include: 1) Comparative studies between Bagley's essentialism and contemporary constructivism to find a balanced curriculum model. 2) Empirical studies on the effectiveness of integrating the core curriculum with 21st-century skills in local contexts. 3) Education policy analysis that assesses the extent to which national curriculum reforms reflect the principles of modern essentialism. 4) Development of a teacher education model that combines Bagley's classical values with digital pedagogical capabilities.

At the policy level, Bagley's ideas can form the basis for educational reform that emphasizes: 1) Equity in quality, not just access. 2) Academic integrity, not just rapid innovation. 3) Social character formation, not just competitive skills.

Thus, William C. Bagley's intellectual legacy holds not only historical significance but also strategic relevance for global education in the era of the 4.0 industrial revolution and the 5.0 industrial revolution.

Funding: This study received no funding.

Conflict of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics Approval: Not applicable.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

- Apple, M. W. (2011). Democratic education in neoliberal and neoconservative times. *International Studies in Sociology of Education*, 21(1), 21–31. <https://doi.org/10.1080/09620214.2011.543850>
- Bagley, W. C. (1905). *The educative process*. Macmillan.
- Bagley, W. C. (1907). *Classroom management*. Macmillan.
- Bagley, W. C. (1920). *Determinants of social education*. Macmillan.
- Bagley, W. C. (1922). *Educational values*. Macmillan.
- Bagley, W. C. (1934). *Education and emergent man*. Macmillan.
- Bagley, W. C. (1939). The Significance of the Essentialist Movement in Educational Theory. *The Classical Journal*, 34(6), 326–344. <http://www.jstor.org/stable/3290875>
- Banks, J. A. (2015). *Cultural diversity and education: Foundations, curriculum, and teaching*. Pearson
- Bates, T. (2024). Teaching the skills needed for the future. In *Creating the University of the Future: A Global View on Future Skills and Future Higher Education* (pp. 123-138). Wiesbaden: Springer Fachmedien Wiesbaden.
- Bestor, A. (1953). *Educational wastelands: The retreat from learning in our public schools*. University of Illinois Press.
- Bestor, A. (1956). *Restoration of learning: A program for redeeming the American school*. Knopf.
- Bowen GA (2009), "Document Analysis as a Qualitative Research Method". *Qualitative Research Journal*, Vol. 9 No. 2 pp. 27–40, doi: <https://doi.org/10.3316/QRJ0902027>
- Brameld, T. (1956). Philosophy of culture: Implications for philosophy of education. *Educational Theory*, 6(3), 158-169. <https://doi.org/10.1111/j.1741-5446.1956.tb01166.x>
- Braun, V., and Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be (com) ing a knowing researcher. *International journal of transgender health*, 24(1), 1-6. <https://doi.org/10.1080/26895269.2022.2129597>
- Cremin, L. A. (1961). *The transformation of the school: Progressivism in American education, 1876-1957*. -Alfred A. Knopf.
- Darling-Hammond, L. (2020). Accountability in Teacher Education. *Action in Teacher Education*, 42(1), 60–71. <https://doi.org/10.1080/01626620.2019.1704464>
- Darling-Hammond, L. (2021). Defining teaching quality around the world. *European Journal of Teacher Education*, 44(3), 295–308. <https://doi.org/10.1080/02619768.2021.1919080>
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., and Osher, D. (2019). 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>
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. Macmillan.
- Dewey, J. (1938). *Experience and education*. Collier Books.
- Duckworth, A. L., and Yeager, D. S. (2015). Measurement Matters: Assessing Personal Qualities Other Than Cognitive Ability for Educational Purposes: Assessing Personal Qualities Other Than Cognitive Ability for Educational Purposes. *Educational Researcher*, 44(4), 237-251. <https://doi.org/10.3102/0013189X15584327>

- Dunn, S. G. (2005). *Philosophical foundations of education: Connecting philosophy to theory and practice*. Pearson/Merrill/Prentice Hall.
- Elias, M. J. (2019). What If the Doors of Every Schoolhouse Opened to Social-Emotional Learning Tomorrow: Reflections on How to Feasibly Scale Up High-Quality SEL. *Educational Psychologist*, 54(3), 233–245. <https://doi.org/10.1080/00461520.2019.1636655>
- Freire, P. (1970). *Pedagogy of the oppressed* (M. B. Ramos, Trans.). Seabury Press.
- Fullan, M., and Quinn, J. (2020). *How Do Disruptive Innovators Prepare Today's Students to Be Tomorrow's Workforce?: Deep Learning: Transforming Systems to Prepare Tomorrow's Citizens*. <https://doi.org/10.18235/0002959>
- Giroux, H. A. (2011). *Education and the Crisis of Public Values: Challenging the Assault on Teachers, Students, and Public Education. Counterpoints: Studies in the Postmodern Theory of Education. Volume 400*. Peter Lang New York.
- Giroux, H. A. (2025). *Teachers as intellectuals: Toward a critical pedagogy of learning*. Bloomsbury Publishing.
- Gutek, G. L. (2011). *Historical and philosophical foundations of education* (5th ed.). Pearson.
- Hamzah, M. R., Mujiwati, Y., Khamdi, I. M., Usman, M. I., and Abidin, M. Z. (2022). Proyek profil pelajar pancasila sebagai penguatan pendidikan karakter pada peserta didik (The Pancasila Student Profile project as a means of strengthening character education for students). *Jurnal jendela pendidikan*, 2(04), 553-559. <https://doi.org/10.57008/jjp.v2i04.309>
- Hirsch, E. D., Jr. (1987). *Cultural literacy: What every American needs to know*. Houghton Mifflin.
- Hirsch, E. D., Jr. (1996). *The schools we need and why we don't have them*. Doubleday.
- Hord, S. M. (Ed.). (2004). *Learning together, leading together: Changing schools through professional learning communities*. Teachers College Press.
- Kincheloe, J. L. (Ed.). (2008). *Knowledge and critical pedagogy: An introduction*. Dordrecht: Springer Netherlands.
- Kirschner, P., Sweller, J., and Clark, R. E. (2006). Why unguided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. *Educational psychologist*, 41(2), 75-86.
- Kliebard, H. M. (1992). *Forging the American curriculum: Essays in curriculum history and theory*. Routledge.
- Knight, J. (2008). *Higher education in turmoil: The changing world of internationalization*. Sense Publishers, 219
- Luckin R (2025), "Nurturing human intelligence in the age of AI: rethinking education for the future". *Development and Learning in Organizations: An International Journal*, Vol. 39 No. 1 pp. 1–4, doi: <https://doi.org/10.1108/DLO-04-2024-0108>
- Lickona, T. (2004). *Character matters: How to help our children develop good judgment, integrity, and other essential virtues*. Simon & Schuster.
- Null, J. W. (2003). An Intellectual Progressive Educator: Toward a Rereading of William Chandler Bagley, 1874–1946. *The Educational Forum*, 67(4), 300–307. <https://doi.org/10.1080/00131720308984578>
- Null, J. W. (2007). William C. Bagley and the Founding of Essentialism: An Untold Story in American Educational History. *Teachers College Record: The Voice of Scholarship in Education*, 109(4), 1013-1055. <https://doi.org/10.1177/016146810710900408> (Original work published 2007)
- OECD. (2019). *OECD learning compass 2030: A Series Of Concept Notes*. OECD Publishing.
- OECD (2026), Navigating Life with Low Literacy and Numeracy: New Results from the 2023 Survey of Adult Skills, OECD Skills Studies, OECD Publishing, Paris, <https://doi.org/10.1787/c198a20f-en>.
- Ornstein, A. C., and Levine, D. U. (2008). *Foundation of education* 10th ed., Houghton Mifflin Company
- Ornstein, A. C., & Hunkins, F. P. (2018). *Curriculum: Foundations, principles, and issues* (7th ed.). Pearson.
- Ozmon, H., and Craver, S. M. (2003). *Philosophical foundations of education*. Pearson.
- Postman, N. (1985). Critical thinking in the electronic era. In *National Forum* (Vol. 65, No. 1, p. 4). Honor Society of Phi Kappa Phi.
- Ravitch, D. (2016). *The death and life of the great American school system: How testing and choice are undermining education*. Basic Books
- Ravitch, D. (2013). *Reign of error: The hoax of the privatization movement and the danger to America's public schools*. Vintage.
- Sahin, M. (2018). Essentialism in Philosophy, Psychology, Education, Social and Scientific Scopes. *Online Submission*, 22(2), 193-204. <https://files.eric.ed.gov/fulltext/ED593579.pdf>
- Safitri, A., Wulandari, D., and Herlambang, Y. T. (2022). Proyek penguatan profil pelajar pancasila: Sebuah orientasi baru pendidikan dalam meningkatkan karakter siswa indonesia (The Pancasila Student Profile Strengthening Project: A New Educational Orientation for Enhancing the Character of Indonesian Students). *Jurnal basicedu*, 6(4), 7076-7086. <https://doi.org/10.31004/basicedu.v6i4.3274>
- Schouten, G., and Brighouse, H. (2015). The relationship between philosophy and evidence in education. *Theory and Research in Education*, 13(1), 5-22. <https://doi.org/10.1177/1477878514562149>
- Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education*, 57(4), 620-631. <https://doi.org/10.1111/ejed.12532>

- Taylor, R. D., Oberle, E., Durlak, J. A., and Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child development*, 88(4), 1156-1171. <https://doi.org/10.1111/cdev.12864>
- Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Jossey-Bass.
- UNESCO. (2021). *Reimagining our futures together: A new social contract for education*. UNESCO Publishing.
- Voogt, J., and Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of Curriculum Studies*, 44(3), 299–321. <https://doi.org/10.1080/00220272.2012.668938>
- Young, K., McNamara, P. M., and Coughlan, B. (2017). Authentic inclusion-utopian thinking? Irish post-primary teachers' perspectives of inclusive education. *Teaching and Teacher Education*, 68, 1-11. <https://doi.org/10.1016/j.tate.2017.07.017>
- Zhao, Y. (2012). *World class learners: Educating creative and entrepreneurial students*. Corwin.