INTELLIGENT TECHNOLOGIES IN EDUCATION

Evaluation of the AI-Teacher Teaching Tasks Spectrum via Practitioner Review

Josiah Koh[1], Michael Cowling[2], Meena Jha[2], Kwong Nui Sim[2]

[1]University of Newcastle, Australia, [2]Central Queensland University, Australia

Abstract

This pilot study aims to evaluate the AI-Teacher Teaching Task Spectrum (AITTTS), a framework designed to categorise human-AI intervention levels based on the teaching tasks and their suitability for AI or human intervention. The primary objective was to provide preliminary validation of the framework's practical utility by examining its alignment with current literature and gathering practitioner feedback. A systematic literature review was conducted, focusing on three key studies that offered insights into AI-teacher task delegation. Additionally, a structured survey was used to collect data from three expert practitioners in AI and education, with Fleiss' Kappa applied to measure agreement. The findings indicated substantial agreement (Fleiss' $\kappa = 0.73$) on the framework's validity, particularly for identifying tasks suitable for AI, such as procedural and knowledge-based activities. However, the study's small sample size and limited geographic diversity restrict the generalisability of the findings. Disagreements, especially regarding AI's role in creative and relational tasks, highlight areas requiring further exploration. As the first step in an iterative research agenda, this pilot study provides foundational insights into the framework's potential. Future research will involve expanded participant samples, broader educational contexts, and iterative refinements to enhance the framework's applicability and generalisability across varied educational settings.

Editors

Section: Educational Technology
Co-Editor-in-Chief: Dr Marios Kremantzis
Associate Editor: Dr Xue Zhou

Publication

Received: 27 August 2024
Revised: 14 February 2025
Accepted: 9 March 2025
Online: 14 March 2025

Keywords

Artificial Intelligence in Education, Artificial Intelligence Pedagogy, AI-Teacher Collaboration, AI-Teacher Teaching Tasks Spectrum

Citation

Koh, J., Cowling, M., Jha, M., & Sim, K. N. (2025). Evaluation of the AI-Teacher Teaching Tasks Spectrum via Practitioner Review. *Intelligent Technologies in Education*, Advanced Online Publication.

Introduction

Artificial intelligence (AI) is increasingly present in education, offering both opportunities and challenges for educators. However, it would be an overstatement to claim that AI has already reshaped the landscape entirely. Its influence, while growing, is still developing in areas such as ethics, finance, and pedagogy. Rather than asserting AI has already revolutionised education, this paper acknowledges that AI holds significant potential to transform teaching and learning as it becomes more integrated into educational environments.

A critical gap exists in the literature: while artificial intelligence (AI) is increasingly adopted in educational settings, there is limited research on how AI and teachers can effectively collaborate across different types of teaching tasks. Understanding AI's role is essential to avoid both underutilisation and overreliance on technology. The AI-Teacher Teaching Task Spectrum (AITTTS) framework aims to address this issue by offering a structured approach to AI-teacher collaboration. This pilot study represents the first step in a broader research agenda designed to evaluate the AITTTS framework. As a pilot validation, it establishes a foundation for future research that will involve iterative refinement through diverse participant samples, expanded educational contexts, and additional rounds of evaluation. This process ultimately aims to provide a robust and generalisable model for AI-teacher collaboration that can be tested through real-world application. To evaluate the practical utility of the AITTTS framework, this study employs a structured survey to gather quantitative and qualitative data from experienced practitioners, along with a systematic literature review, to evaluate the AITTTS framework in practical settings. This study was framed as a pilot to explore the practical utility of the AITTTS framework in educational settings. The focus on a small sample and targeted literature review was intentional, allowing for a manageable scope and providing preliminary insights to guide future research. Practitioner reviews are particularly valuable in contextualising frameworks like the AITTTS within everyday teaching practices, ensuring that the framework meets the needs of educators and aligns with classroom dynamics (Alharmoodi & Lakulu, 2022; Bradford et al., 2019) and grounded in practical experience.

This study seeks to answer the following research questions:

RQ1: How does the AI-Teacher Teaching Task Spectrum fit within the broader literature on AI's role in education?

RQ2: Is there sufficient consensus among independent practitioners regarding the validity and practical utility of the AI-Teacher Teaching Task Spectrum?

Clarifying these questions will provide a foundation for refining the AITTTS framework and, more broadly, for better understanding how AI and human teachers can collaborate effectively in educational settings.

The AITTTS framework draws on Bloom's Taxonomy (1985), which categorises teaching tasks from procedural duties, such as administrative responsibilities, to more complex tasks like creativity and pastoral care. By mapping AI capabilities against this taxonomy, the AITTTS framework offers a method to determine where AI can support or enhance teaching, and where human intervention remains crucial. Evaluating the effectiveness of this framework is essential to ensure that AI is integrated in ways that enhance, rather than diminish, the vital role of educators.

Given the nascent stage of research on the AI-Teacher Teaching Task Spectrum (AITTTS) and the limited geographic and participant diversity, this study is framed as a pilot. The aim is to provide preliminary insights into the framework's utility and identify directions for future, larger-scale research.

Background

Artificial intelligence (AI) in education has traditionally been used to handle routine, repetitive tasks that reduce the administrative burden on teachers. Early AI applications primarily focused on systems designed to automate administrative tasks such as grading assignments, managing student records, and responding to frequently asked questions (Shen & Su, 2020). Tools like automated grading platforms used algorithms to evaluate objective assessments, providing rapid feedback to students, particularly for multiple-choice questions and structured essays. Chatbots and virtual assistants were commonly employed to handle student queries related to deadlines, course information, and institutional policies, effectively streamlining these operations and allowing educators to focus more on pedagogical aspects of teaching.

AI-driven learning management systems (LMS) also played a pivotal role in this phase, automating tasks such as tracking student performance, managing submissions, and delivering instructional content in a more structured way (Pérez et al., 2020). These early AI tools demonstrated the potential of AI to assist with the procedural and administrative tasks of education, reducing the time teachers spent on non-pedagogical duties and enabling them to dedicate more time to personalised instruction.

Expansion into Instructional and Learning Tasks

As AI technologies have advanced, newer systems have expanded beyond basic administrative functions to take on more complex instructional roles. Intelligent tutoring systems (ITS), for instance, are designed to provide students with personalised learning experiences by adapting to their learning needs in real-time (Ashri & Sahoo, 2021). These systems use machine learning algorithms to track student progress, identify areas where they need improvement, and adjust the difficulty of the tasks accordingly. This allows for a more student-centred approach, where AI plays a more active role in knowledge recall and basic analysis tasks—corresponding to the foundational levels of Bloom's Taxonomy.

Automated essay scoring systems have also emerged, providing feedback on more subjective assessments, though their accuracy and reliability remain under scrutiny, particularly when it comes to assessing creativity and nuance in student work (Williamson et al., 2020). AI-powered adaptive learning platforms, like those used by online education providers, further push the boundaries by offering real-time feedback and dynamically adjusting content to suit individual student learning paths.

AI's Growing Role in Higher-Order Teaching Tasks

While AI has proven effective in assisting with knowledge recall and basic instructional tasks, its application to higher-order teaching tasks, such as critical thinking, creativity, and relational dynamics, remains more limited. AI tools like predictive analytics platforms and real-time feedback systems can help provide data-driven insights into student performance, but the teacher's role is still critical in guiding students through the process of contextualising information, applying it to real-world problems, and drawing informed conclusions.

For example, while AI-driven systems can assist with generating data for analysis, teachers are still needed to help students critically interpret that data, fostering deeper engagement with complex topics. In these tasks, AI serves as a supportive tool, while the human teacher plays the lead role in encouraging higher-order thinking and creative problem-solving (Cotton et al., 2023).

At the upper end of the AI-Teacher Task Spectrum, tasks that require emotional intelligence and creativity—such as pastoral care and motivating students—remain largely outside the reach of current AI technologies. AI tools, while useful for providing supportive content or handling procedural duties, lack the emotional depth and relational capabilities that are essential in teaching tasks where empathy, motivation, and inspiration play a critical role (Young, 2022). AI chatbots could also not be able to recognise or appropriately respond to signs of distress or possible harm (De Freitas et al., 2023). In this context, the human teacher is irreplaceable, offering the personal connection that is central to a holistic educational experience (Selwyn, 2019).

Literature

AI's early applications in education focused primarily on simplifying administrative tasks, but recent advancements have extended its reach into more complex areas of teaching and learning. This shift necessitates a deeper understanding of the role AI plays in education and how AI and human teachers can collaborate most effectively. Such an understanding also provides a foundation for examining the AI-Teacher Teaching Task Spectrum (AITTTS)—a framework designed to navigate these evolving roles by delineating tasks that can be delegated to AI and those that require the unique capabilities of human teachers.

Early Roles of AI in Education

The initial role of AI in education was primarily focused on automating routine, repetitive tasks to reduce the administrative burden on teachers. Early AI systems were designed to handle tasks such as grading assignments, managing student records, and providing automated responses to frequently asked questions (Pérez et al., 2020). These rule-based systems, including AI-powered grading platforms, used algorithms to evaluate student responses to multiple-choice questions and structured essays. Automated systems like these freed teachers from time-consuming grading processes, allowing them to focus more on student engagement and higher-order tasks such as critical thinking and personalised feedback (Luckin et al., 2016).

AI as an Administrative Tool

In this phase, AI was largely confined to administrative and procedural roles, functioning as an assistant rather than a central player in the educational process. For example, automated grading tools provided rapid feedback to students on objective assessments, significantly speeding up the evaluation process. Similarly, chatbots and virtual assistants were introduced in many institutions to answer routine student queries, such as questions about deadlines, course materials, or institutional policies (Shen & Su, 2020). These AI-driven systems made it easier for students to access information and allowed teachers to redirect their energy toward more complex pedagogical tasks.

Impact on Education

The impact of these early AI implementations was twofold: while AI relieved teachers of repetitive tasks, it also began to reshape the classroom dynamic. The time saved through automated systems allowed for more personalised instruction, as teachers could now devote additional attention to individual student needs (Pérez et al., 2020). Studies found that this shift improved student engagement and led to more efficient classroom management. By automating tasks such as grading, teachers were also able to offer students more timely feedback, helping to foster a more responsive and dynamic learning environment (Chen et al., 2020).

Concerns about AI's Early Use in Education

However, the early adoption of AI in education was not without its challenges. One of the main concerns was the over-reliance on AI for grading, especially for subjective assessments such as essays and projects, where human judgment is often necessary to evaluate creativity, nuance, and context. Critics argued that AI-based grading systems could not account for the complexity of human expression and were prone to errors in interpretation (Williamson et al., 2020). Additionally, the feedback provided by these systems was sometimes perceived as too simplistic or impersonal, lacking the depth and guidance that a teacher's input could offer (Wang et al., 2024).

Another concern was the quality and reliability of the AI systems themselves. As these technologies were still in their early stages, many systems faced software glitches and operational limitations. For example, AI-driven feedback systems, while fast, sometimes delivered feedback that was not useful or relevant to the specific context of the student's learning (Wang et al., 2024). This led to frustrations among both students and educators, who expected more meaningful insights from the technology. Moreover, there were ethical concerns about data privacy, as AI systems in educational settings collected significant amounts of personal data from students, raising questions about how that data was being used and stored (Holmes et al., 2019).

Balancing Efficiency with Pedagogy

While AI brought greater efficiency to the educational process, its impact on pedagogy was more complex. Teachers and educational institutions had to find a balance between the benefits of automation and the importance of maintaining the human touch in education. For example, while AI could handle objective tasks such as grading multiple-choice tests, teachers were still needed to provide deeper insight and understanding, particularly when it came to guiding students through more subjective or creative work. The human element remained critical for tasks requiring emotional intelligence, creativity, and critical thinking (Selwyn, 2019).

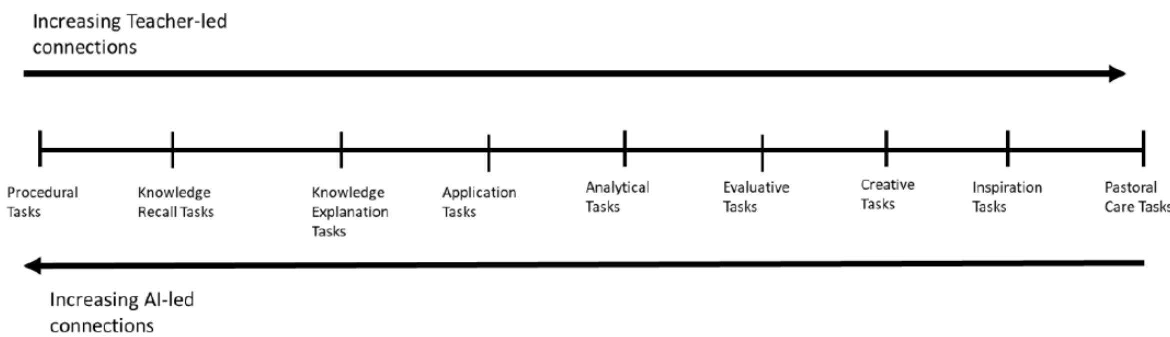
As AI began to take on more administrative tasks, concerns also grew about the potential for teachers to lose some of the relational aspects of teaching. AI’s ability to streamline processes sometimes led to a depersonalisation of education, as interactions between teachers and students became more transactional. While this efficiency was beneficial, it raised questions about how AI might affect the relational dynamics in education, a critical component for student motivation and engagement.

AI-Teacher Teaching Tasks Spectrum

The AI-Teacher Teaching Task Spectrum (AITTTS) is a conceptual framework designed to help delineate the appropriate roles for AI and human teachers within the educational process. As AI systems have evolved and become capable of performing many tasks traditionally handled by educators, there has been a growing need for clarity about where AI can excel and where human intervention remains indispensable. The AITTTS (Figure 1) addresses this by mapping teaching tasks along a spectrum, from procedural and administrative duties to creative, relational, and pastoral care tasks.

Figure 1

Spectrum of AI-Teacher Teaching Tasks



At the lower end of the spectrum are procedural tasks, which can largely be delegated to AI systems. These include administrative responsibilities such as managing assessment deadline extensions, handling basic queries, and guiding students to course information. AI systems, such as those integrated into learning management systems, have already been shown to effectively manage these tasks (Pérez et al., 2020). Their primary advantage lies in reducing the time teachers spend on repetitive duties, allowing them to focus more on personalised instruction and student engagement.

The next level includes knowledge recall tasks, which align with the foundational levels of Bloom's Taxonomy. AI-driven tools like intelligent tutoring systems are particularly suited to assisting students with recalling information and performing basic analysis. These systems can access vast amounts of information instantaneously, serving as knowledge repositories that students can use to enhance their learning experience (Ashri & Sahoo, 2021). However, while AI can provide students with access to information, the teacher's role is critical in guiding them to apply and interpret this knowledge in meaningful ways.

As the tasks become more complex—such as application, analysis, and evaluation—the role of human teachers becomes more prominent. AI can support these tasks by providing initial data or performing basic analyses, but higher-order thinking tasks often require human oversight to ensure that students are not only using AI tools responsibly but also interpreting the data critically (Cotton et al., 2023). For example, AI might help a student retrieve and organise relevant data, but it is the teacher who guides the student in contextualising that data, applying it to real-world problems, and drawing informed conclusions.

At the upper end of the spectrum are creative and relational tasks, where the human element of teaching is essential. Tasks that require emotional intelligence, empathy, and creativity are inherently human and cannot be effectively handled by AI. Teachers play a crucial role in fostering creativity in students by helping them synthesise knowledge, apply it in innovative ways, and navigate the emotional and relational aspects of learning. AI's role in these tasks is minimal, limited to providing tools or resources, while teachers guide the creative process and offer personal support.

Finally, teaching is not merely an intellectual exercise but also a relational one. Inspirational tasks, such as motivating and encouraging students, and pastoral care tasks, which provide emotional and personal support, are deeply connected to the teacher-student relationship. While AI can simulate some aspects of human interaction, it lacks the emotional depth and lived experiences that are central to genuine relational teaching (Williamson et al., 2020). In these areas, the human teacher remains irreplaceable, offering empathy, inspiration, and a personalised approach to each student's needs.

Method

Research Process

This pilot study employed a three-phase research process to provide an initial evaluation of the AI-Teacher Teaching Task Spectrum (AITTTS). Given the exploratory nature of this work, the study was designed to test the framework's feasibility and gather preliminary insights from a small sample of expert practitioners. The process involved a systematic literature review, practitioner evaluation through a semi-structured questionnaire, and a final phase of data analysis. These methods were selected to rigorously assess the validity and reliability of the framework and to provide a comprehensive evaluation through both theoretical and practical lenses (Hasanpoor et al., 2019).

The study combined these approaches to provide a holistic understanding of the AITTTS. The literature review established the theoretical foundation for understanding AI's role in education by identifying key trends and existing frameworks. This review addressed RQ1, which seeks to understand how the AITTTS fits within the broader literature on AI in education. In contrast, the empirical phase through practitioner feedback directly addressed RQ2 by testing the framework's practical validity and exploring practitioner consensus on its utility in real-world settings.

Theoretical Framework

The AITTTS is theoretically grounded in Bloom's Taxonomy and an adapted Technology Acceptance Model (TAM). Bloom's Taxonomy categorises educational objectives into a hierarchy of cognitive skills, ranging from simple recall to complex evaluation and creation tasks, which directly informs the classification of tasks within the AITTTS. This theoretical model helps distinguish tasks that AI can take on, such as procedural or knowledge-recall tasks, from those that require human intervention, like creative or pastoral care tasks.

The adapted Technology Acceptance Model (TAM) provides additional insight into the relationship between AI and human teachers by assessing the perceived usefulness and ease of use of AI tools in educational contexts. This hybrid approach, adapted to the educational context, emphasises the role of social, cognitive, and teaching presence, offering a nuanced understanding of how AI can support human teachers without fully replacing them.

Phase One: Systematic Literature Review

The first phase of the study was a systematic literature review designed to identify existing knowledge related to the role of AI in education, specifically focusing on the interaction between AI and teachers. A systematic review was selected to ensure a transparent and replicable methodology, following the typology outlined by Grant & Booth (2009). This review allowed for the contextualisation of the AITTTS within the broader literature and helped to define the theoretical foundations of the framework. The systematic literature review was designed to establish a theoretical foundation for this pilot study by identifying key trends and frameworks relevant to AI-teacher interactions. While the review captured the state of the field up to late 2023, the focus was deliberately narrowed to align with the study's exploratory scope, selecting studies that offered direct insights into task delegation. The primary objectives were:

1. To identify key themes and concepts regarding the role of AI in education.
2. To explore any gaps, inconsistencies, or controversies within the existing literature.
3. To position the AITTTS within the current educational discourse.

Phase Two: Practitioner Review

The second phase involved a practitioner review as part of this pilot study, in which three experienced educators and teaching professionals with AI exposure were engaged to evaluate the validity of the AITTTS. A structured survey was employed, comprising closed-ended questions with Likert-scale responses to facilitate quantitative analysis, and open-ended questions to capture qualitative insights. Although the survey was initially described as semi-structured, its implementation leaned toward a structured format, limiting participants' ability to deviate from predefined responses. This design choice was intentional to ensure consistency across responses within the constraints of the small sample size, while the qualitative questions allowed for richer, contextual reflections. Combining quantitative and qualitative methods in this pilot study was deemed appropriate for generating preliminary insights into the framework’s utility, acknowledging that the small sample size limits the generalisability of findings. . The key steps were:

1. Selecting a panel of experienced practitioners with AI exposure in their teaching.
2. Administering a semi-structured questionnaire designed to assess the framework’s value and alignment with practical experiences.
3. Evaluating the feedback for alignment with the AITTTS framework, focusing on whether the framework accurately reflects the teaching tasks in contemporary educational settings.

This empirical phase was conducted to complement the theoretical findings from the literature review, providing a well-rounded evaluation of the framework.

Final Phase: Data Analysis

As part of this pilot study, the data from both phases underwent rigorous analysis tailored to the pilot study's exploratory goals. Quantitative responses were analysed using Fleiss' Kappa to assess agreement among the three practitioners, while qualitative responses were analysed thematically to identify emerging patterns and contextual insights. Although the small sample size limits statistical power, this mixed-methods approach was chosen to balance measurable practitioner consensus with nuanced feedback, providing a holistic preliminary evaluation of the AITTTS framework. This approach, while limited in scope, provides an initial foundation for evaluating the AITTTS framework. Thematic analysis was chosen for its adaptability in identifying patterns and trends from qualitative data (Nowell et al., 2017), while Fleiss' Kappa was applied to measure consensus on the validity of the AITTTS.

Results

Phase One Results: Systematic Literature Review

A systematic literature review was conducted to establish a theoretical foundation for the AI-Teacher Teaching Task Spectrum (AITTTS). An initial search yielded 86 records, which were filtered based on relevance to the research objectives. Studies were included if they explicitly addressed AI-teacher task delegation or provided a conceptual framework aligned with the AITTTS. Fourteen studies met these criteria, of which three were selected for detailed discussion due to their direct alignment with the framework's focus on task delegation between AI and human teachers. The remaining studies provided broader contextual insights and were summarised in the background section.

Search strategy

The literature review was conducted using Scopus, selected due to its comprehensive coverage of peer-reviewed literature in the fields of education and AI (Aghaei Chadegani et al., 2013). The review focused on articles published between 2016 and 2024, capturing the most recent advancements in AI education, particularly after the development of major AI tools such as ChatGPT. The other inclusion criteria were: (1) articles published in English, (2) peer-reviewed journal articles or conference proceedings, and (3) studies that explicitly addressed AI-teacher interactions or AI applications in education. Articles that focused solely on technical aspects of AI without relevance to education were excluded.

Table 1

List of Search Terms

Topics	Search Term
Artificial Intelligence	*"AIed*"* OR *"Artificial Intelligence in education*"* OR *"Teacherbot*"* OR *"automated tutor*"* OR *"intelligent computer-aided instruction"*
AND Teacher	"Teach*" OR "Instruct*" OR "Tutor*"
AND Framework	"Frame*" OR "Paradigm*" OR "Princ*"

After the application of the search terms in Table 1, a total of 86 records were identified. This set of records were useful as a frame of reference but was narrowed down for clarity and recency. A second set of criteria was used to sharpen the focus of the search. The inclusion criteria are seen in Table 2 below. Using records in the past 8 years would provide the most recent and updated literature in this space. Given that OpenAI (the creators of ChatGPT) was only founded in December 2015, going back to 2016 will give a long enough timeframe to see any proposed frameworks or paradigms for AI in education, but also recent enough so as to have relevance and currency.

Table 2

Inclusion Criteria

Inclusion Criteria	Time Period: from 2016 to 2024
	Publication Type: Peer Reviewed Journals, Conference Proceedings

	English

A total of 49 records were found to fit the new requirements. Out of these 49 articles, articles that proposed a framework but was narrowly focused only on the study were excluded. This is because the research is trying to have a broader view of AI in education, and how the AI-Teacher teaching tasks spectrum can sit within it. After reading through manually and checking for relevancy, a final total of fourteen articles are found.

Table 3

List of Relevant Articles

Authours (Year)	**Title of article**	**Framework / Paradigm**
Ouyang F.; Jiao P.; Alavi A.H. (2020)	Artificial intelligence-based smart engineering education	leaner-receiver, learner-partner and learner-centre
Lameras, Petros (2022)	A Vision of Teaching and Learning with AI	Content and information, AIED and Knowledge application, AIED and adaptive tasks, AIED and adaptive assessment, AIED and self-regulation
Lameras P.; Arnab S. (2022)	Power to the Teachers: An Exploratory Review on Artificial Intelligence in Education	Adaptive, Collaborative Learning Support (ACLS), Conversation and Social and Emotional Learning

Baker, T.; Smith, L.; Anissa, N. (2019)	Educ-AI-Tion Rebooted? Exploring the future of artificial intelligence in schools and colleges	(1) learner-facing, (2) teacher-facing, and (3) system-facing
Xu, W & Ouyang, F (2022)	A systematic review of AI role in the educational system based on a proposed conceptual framework	complex adaptive systems perspective
Hwang G.-J.; Xie H.; Wah B.W.; Gašević D. (2020)	Vision, challenges, roles and research issues of Artificial Intelligence in Education	Intelligent tutor, Intelligent tutee, Intelligent learning tool or partner, Policy-making advisor
Alam A. (2021)	Should Robots Replace Teachers? Mobilisation of AI and Learning Analytics in Education	AI Education Administration, AI Instruction, AI Learning, AI performance
Sottilare R.A.; Baker R.S.; Graesser A.C.; Lester J.C. (2018)	Special Issue on the Generalized Intelligent Framework for Tutoring (GIFT): Creating a Stable and Flexible Platform for	Authoring, instructional management, evaluation

	Innovations in AIED Research	
Gibson D.; Kovanovic V.; Ifenthaler D.; Dexter S.; Feng S. (2023)	Learning theories for artificial intelligence promoting learning processes	Macro, micro, meso levels of learning
Sowmia K.R.; Poonkuzhali S. (2020)	Artificial intelligence in the field of education: A systematic study of artificial intelligence impact on safe teaching learning process with digital technology	Smart Coaching, Evaluation and Appraisal, framework and personalisation, preparing and forecasting
Feng S.; Law N. (2021)	Mapping Artificial Intelligence in Education Research: a Network-based Keyword Analysis	online learning; (2) game-based learning; (3) collaborative learning; (4) assessment; (5) affect; (6) engagement; and (7) learning design
Ouyang, F & Jiao, P. (2021)	Artificial intelligence in education: The three paradigms	AI-Empowered, AI-supported, AI-Directed

Humble, N & Mozelius, P (2019)	Teacher-supported AI or AI-supported teachers?	Teacher-supported AI, AI-Compatible AI, AI-Supported Teachers
Lodge, J. M; Thompson, K & Corrin, L (2023)	Mapping out a research agenda for generative artificial intelligence in tertiary education	AI in Assessment Design & Integrity, AI in Personalised and Adaptive Learning, AI in Teacher Workload Support, Critical Thinking and AI Literacy

The systematic literature review was conducted in late 2023 to capture studies relevant to the evolving field of AI in education. The search focused on publications from 2016 to 2023, a period chosen for its relevance to recent advancements in AI technologies, particularly following the introduction of tools like ChatGPT. While the search yielded 86 records, these results represent the state of the field at the time of review. Subsequent replications of the search query in early 2024 have indicated a substantial increase in publications, reflecting the rapid growth of research in this area. The original search strategy was deliberately refined to focus on studies directly aligned with the research question. Criteria such as relevance to AI-teacher interactions and the presence of conceptual frameworks guided the selection process, resulting in 14 highly relevant studies for detailed analysis.

Role of the AI in the context of the AI-Teacher Relationship

Out of the 14 studies identified as relevant, three were prioritised for detailed discussion due to their direct focus on AI-teacher task delegation, a core element of the AITTTS framework. These studies—Ouyang & Jiao (2021), Humble & Mozelius (2019), and Lodge et al. (2023)—provided the most pertinent insights into the delineation of tasks between AI and human teachers. The remaining studies offered valuable but broader perspectives on AI in education, which were summarised in the background section to contextualise the findings.

Ouyang and Jiao's (2021) paradigm

Ouyang and Jiao's (2021) work explored AI in education across three paradigms: AI-directed, AI-supported, and AI-empowered.

AI-directed Paradigm: This paradigm takes a more behaviorist approach, viewing AI as the best source of domain information or knowledge. As such, in this perspective, students are static recipients, being directed down predetermined learning pathways, with the AI acting as the 'sage on the stage'. AI's focus in this paradigm is mainly centred on delivery and operation, and less on the adaptation of student needs. This paradigm is AI-led and less learner-centric.

AI-supported Paradigm: The AI-supported paradigm is on the other end of the ideological spectrum. Embracing a cognitive, social constructivist approach, this paradigm treats students and AI as co-collaborators in learning delivery. AI adapts to students needs and requests, which in turn shapes how students respond to AI, moving in this Ouroboros cycle of learning and adaptation. Ouyang and Jiao (2021) argue that this paradigm supports models that require cognitively higher order tasks, such as analytical thinking and creative problem solving.

AI-empowered Paradigm: The AI-empowered paradigm is the 'middle of the road' view. Rooted in a connectivist view, this user-led paradigm views AI as an educational tool, rather than as a collaborator or as the best source of knowledge. It posits that human teachers that the lead in education, whilst advanced AI tools will support and follow these human teachers, providing real-time knowledge about students for just-in-time interventions or teaching. Combining the human touch and AI, this paradigm believes that students will feel self-empowered leading to a greater self-ownership of learning and expanding knowledge horizon.

Humble and Mozelius categories

Humble & Mozelius (2019) categorised the relationship between AI and human teachers into three main types: AI-supported Teacher, Teacher-Supported AI, and Teacher-Compatible AI.

AI-supported Teacher: The AI-supported teacher is one where the teacher still serves as the 'sage on the stage' but supported by AI as an information provider. In this format, teaching pedagogy still primarily follows the traditional mode of teaching, where the teacher is provided with AI analysed student analytics, and AI can support the curriculum & assessment design and to some extent, delivery.

Teacher-Supported AI: The teacher supported AI is the inverse of the AI-supported teacher. Here, the tasks are split between AI and the teacher, with the AI dictating most of the learning paths and content delivery. Teachers will support the learning by prompting and encouraging critical thinking and application. In this category, the AI will almost always start the initial learning journey, and the human teacher will then come in to fill in the other aspects of learning.

Teacher-Compatible AI: Similar to the AI-empowered paradigm above, the teacher-compatible AI is a more 'middle ground' category, but the teacher is still primarily the leading authourity in the teaching tasks. AI acts as a teacher aide, and will acts on tasks delegated by the teacher, following the teacher's direction but with limited autonomy. However, clearly distinguishing between teacher-supported AI and AI-supported teachers is very difficult, hence a teacher-compatible AI, combining human and AI strengths, could be a good solution that incorporates the best of both worlds.

Lodge's research agenda

In the Australian Journal of Education Technology (AJET)'s 2023 editorial, Lodge and colleagues (2023) mapped out a research agenda for AI, looking more specifically at assessment integrity, assessment redesign, and learning and teaching. However, the mapping has also indirectly highlighted the roles of AI and where it could sit on the various paradigms and on the AITTTS.

AI could challenge the view of what constitutes cheating and academic misconduct. Viewing AI in the lens of AI being a writing tool or referencing source would mean that assessments can be written and understood differently. Instead of assessing for knowledge recall and explanation, from an AITTTS perspective, assessments could assess for higher order cognitive skills. This would also mean re-interpreting what assessment and academic integrity looks like. AI will now be allowed to be used under certain conditions, and form part of the assessment rubric.

Given that academic and assessment integrity is viewed in a different light, assessment redesign will naturally also change. Lodge proposes a substantial shift from shunting AI to embracing AI, to improve on current known weaknesses in assessment design (such as a lack of individualised testing) whilst acknowledging the potential of AI to further refine assessment. AI enabled assessments could leverage on AI for additional support and real-time feedback, allowing for a more conversational assessment style, encouraging students to apply critical thinking to AI-generated outputs.

The most controversial and contentious piece of the puzzle is the learning and teaching with AI piece. Although AI's effectiveness in giving procedural feedback has been documented in early studies, its long-term effectiveness as an AI tutor is unknown. Adding to this complexity is that education is crucially a relational endeavour, where interpersonal relationships inspire, motivate and encourage creativity. Holmes and colleagues (2023) further explore AI in education's possibilities, suggesting different forms of teaching and learning, such as collaborative learning, continuous assessments, AI learning companions, and teaching tools.

Literature Gap

While Ouyang and Jiao (2021) categorise AI in education into AI-directed, AI-supported, and AI-empowered paradigms, these models primarily focus on the technical capabilities of AI rather than the pedagogical implications of AI-Teacher interactions. Their framework, though comprehensive, lacks a clear integration of relational and pastoral aspects of teaching, which are critical in holistic education. Similarly, Humble and Mozelius (2019) provide a typology of AI-teacher relationships, yet their categorisation does not sufficiently address the dynamic interplay between AI and teachers across the spectrum of teaching tasks, particularly in areas requiring creativity and emotional intelligence.

This gap in the existing frameworks underscores the need for a more nuanced model that not only categorises AI’s technical roles but also aligns these roles with the pedagogical and relational dimensions of teaching. The AI-Teacher Teaching Tasks Spectrum (AITTTS) proposed in this study seeks to fill this gap by offering a more holistic view that integrates both the cognitive and emotional aspects of teaching. By situating the AITTTS within the broader context of existing AI education frameworks, this study aims to provide a more comprehensive tool that educators can use to navigate the complexities of AI integration in teaching.

Phase Two Results: Practitioner Review

In Phase 2, a pilot practitioner review was conducted to further evaluate the AI-Teacher Teaching Task Spectrum (AITTTS) framework, building on the insights from the literature review in Phase 1. The primary goal of this phase was to gather feedback from active educational practitioners, combining their hands-on experience with their theoretical expertise in AI integration. A structured survey was distributed to practitioners who have experience using AI tools in educational settings, ensuring a balanced collection of both quantitative and qualitative data.

Qualitative responses revealed three key themes: (1) AI's strong suitability for procedural tasks, where participants highlighted time-saving benefits; (2) challenges in integrating AI into creative and relational tasks, with concerns about AI's lack of emotional intelligence and context sensitivity; and (3) the need for teacher training to effectively leverage AI. For instance, one participant noted, 'AI excels in routine tasks like grading, but it cannot yet replicate the human touch required for inspiring or motivating students. The survey instrument comprised both closed-ended and open-ended questions, designed to evaluate the AITTTS framework across procedural, instructional, and creative teaching tasks. Closed-ended questions used a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree) to measure participant agreement with task classifications. Open-ended questions invited practitioners to elaborate on the suitability of AI for specific tasks and their experiences with AI integration in teaching.

To analyse the quantitative responses, Fleiss Kappa was applied to measure the level of agreement among participants. Although the sample size (N=3) was small, this phase was designed as a pilot study, aimed at collecting preliminary insights on the framework's practical utility. The findings from this initial phase offer valuable direction for future research, which will involve larger and more diverse samples to validate the results and ensure greater generalisability.

The survey instrument was developed based on prior literature on AI's role in education, ensuring that the questions aligned with the key objectives of the AITTTS framework. By engaging practitioners in this review, the study provides an important first step in testing the AITTTS framework's validity, relevance, and utility in real-world educational settings. Despite the limited sample size, the combination of quantitative agreement and qualitative insights offers a robust initial test of the framework. The survey consisted of closed-ended questions designed to evaluate the AITTTS framework across multiple categories of teaching tasks—procedural, instructional, and creative. These questions used a Likert scale to measure participants' agreement with the classification of tasks. Additionally, open-ended questions were included to gather qualitative insights, allowing practitioners to provide deeper reflections on their practical experience with AI in teaching. This mixed-methods approach ensured that the study captured both the level of consensus and the nuanced perspectives of each participant.

The consensus from these practitioners is invaluable for assessing the practical robustness of the AITTTS framework. Their feedback also helps evaluate how well the framework aligns with current best practices and its adaptability to future developments in AI-assisted teaching.

Procedure

A panel of practitioners was selected through purposive sampling. Three academic staff members from a vocational and tertiary educational institution in New Zealand were independently chosen for their experience and expertise in both education and AI integration. The practitioners had varying levels of experience with AI tools, ensuring a broad range of insights:

- **Participant 1,** a senior academic staff member with 10 years of teaching experience, regularly used AI tools like ChatGPT for learning design and tutoring support.

- **Participant 2**, another academic staff member with over 10 years of teaching, had experimented with AI chatbots (primarily ChatGPT) in student interactions.
- **Participant 3**, with 12 years of teaching experience, had primarily used AI in his research (across various AI research tools like Julius AI and ChatGPT) but had not yet incorporated it into his teaching practice.

The sample size of three practitioners was chosen to align with the exploratory nature of this pilot study, balancing resource and time constraints while ensuring informed and relevant feedback for the AITTTS framework. Each participant brought distinct expertise: while two practitioners had direct experience with AI integration in teaching, the third participant's expertise in AI research offered complementary insights into the broader applicability of AI in education. Prior to data collection, participants received comprehensive information sheets detailing the study's objectives, response instructions, and confidentiality assurances. The researcher maintained availability throughout the process to address queries, and all responses were collected within 48 hours without the need for reminders. While this sample size limits the generalizability of the findings, it provides a valuable foundation for understanding the framework's utility in practice.

Table 4

Demographics of practitioners

	Years in Education	Position	Gender
Practitioner 1	10	Senior Academic Staff Member	Male
Practitioner 2	10+	Academic Staff Member	Male
Practitioner 3	12	Academic Staff Member	Male

This table provides a demographic breakdown of the practitioners, including their years in education, positions, and gender.

The study was conducted in compliance with the ethical guidelines set by the research office of the vocational institution. The survey instrument comprised both closed-ended and open-ended questions, designed to evaluate the AITTTS framework across procedural, instructional, and creative teaching tasks. Closed-ended questions used a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree) to measure participant agreement with task classifications. Open-ended questions invited practitioners to elaborate on the suitability of AI for specific tasks and their experiences with AI integration in teaching. The survey design was informed by foundational studies on AI in education, including Pérez et al. (2020) and Shen & Su (2020), ensuring alignment with established literature. Informed consent was obtained from all participants before the survey was administered. The participants were informed of the study's purpose, and the confidentiality of the collected data was ensured throughout the research process.

Phase Three Results: Data Analysis

The decision to adopt a quantitative approach for phase two was driven by the need to measure consensus among practitioners regarding the AITTTS framework. Fleiss' Kappa was used to quantify the level of agreement between the participants, allowing for a structured analysis despite the small sample size. Although only three practitioners participated in this initial phase, they were selected for their extensive experience with AI in education, ensuring that their feedback was both informed and insightful. While the small sample size is a limitation, early-stage validation studies often begin with a limited number of expert participants to assess the framework's feasibility before broader testing (Hennink & Kaiser, 2022). Future research will expand the participant pool to strengthen the robustness and generalisability of the findings.

For the data analysis, RStudio software was used to calculate Fleiss' κ, which measures inter-rater agreement among three or more raters for categorical data. Fleiss' κ ranges from -1 to +1, where -1 indicates no agreement, 0 suggests agreement equal to chance, and +1 represents perfect agreement. Given that this was a Likert scale survey with five options, a fixed marginal kappa was applied. The survey results are presented below (Table 5).

Table 5

Table of responses

	Strongly Disagree	**Disagree**	**Neither Agree Nor Disagree**	**Agree**	**Strongly Agree**
Procedural Tasks	0	0	0	0	3
Knowledge Recall Tasks	0	0	0	3	0

Knowledge Explanation Tasks	0	0	0	3	0
Application Tasks	0	0	3	0	0
Evaluative Tasks	0	0	3	0	0
Creative Tasks	0	0	2	0	1
Inspirational Tasks	0	0	1	2	0
Pastoral Care Tasks	0	0	0	3	0

Based on the table of responses, the percent overall agreement is 83.33%. The fixed-marginal kappa was calculated to be 0.73. At a 95% confidence interval, it will range between (0.41 to 1.00). The statistical significance threshold for the results is set at $p < 0.05$, with a 95% confidence interval for all cases

Discussion

The AI-Teacher Teaching Task Spectrum (AITTTS) offers a structured framework for understanding the evolving collaboration between AI and human teachers. While the framework aligns with key principles in the existing literature, it is essential to explore both its congruence with established theories and its divergences, particularly in tasks requiring creativity and emotional intelligence. Additionally, the framework's iterative validation process requires further elaboration and substantiation.

Critical Engagement with the Literature

The AITTTS framework aligns well with existing paradigms of AI's role in education, particularly in tasks that involve procedural and administrative duties. For example, frameworks such as Ouyang and Jiao's AI-directed tasks and Humble and Mozelius' teacher-assisted AI support the notion that AI excels in automating routine tasks, such as grading and knowledge recall. The AITTTS's placement of these tasks on the lower end of the spectrum fits within the broader literature, which acknowledges AI’s efficiency in handling repetitive, non-pedagogical duties (Pérez et al., 2020; Shen & Su, 2020).

However, the AITTTS diverges from the existing literature when applied to more complex tasks requiring human intervention, such as creative problem-solving, emotional intelligence, and pastoral care. The current literature consistently points to the limitations of AI in these areas (Cotton et al., 2023; Williamson et al., 2020), and while AI can support tasks like data analysis, it struggles to replicate the nuanced human insights necessary for creativity and relational dynamics. The AITTTS acknowledges these limitations by placing these higher-order tasks firmly in the domain of human teachers, a distinction that reinforces the irreplaceable role of teachers in these areas.

Although this alignment is apparent, it would be an oversimplification to claim that the AITTTS “sits almost perfectly within the broader literature.” While the framework complements existing studies, it also highlights gaps in our understanding of AI's role in higher-order educational tasks. These divergences point to an evolving landscape in which AI's capabilities are rapidly developing but still require significant human oversight. Future research should explore these complexities in greater depth, especially as AI continues to evolve and expand into more nuanced educational tasks. Thus, the AITTTS framework not only aligns with but also challenges the current literature by offering a pathway for integrating AI without diminishing the critical, human-centric aspects of teaching.

Iterative Validation Process

This pilot study serves as an initial validation of the AITTTS framework, examining its feasibility and alignment with practitioner insights. While the findings offer a preliminary evaluation through practitioner reviews, it is important to recognise that framework validation is not a one-time event but rather an ongoing process that must be tested across various contexts and educational environments (Fernández-Gómez et al., 2019). Future iterations of this research will incorporate larger, more diverse participant samples, expanded geographic contexts, and refined survey methodologies to address the complexities of AI-teacher collaboration across varied educational settings. These steps represent early stages in a longer, iterative journey toward comprehensive validation of the framework. While the small sample size provided valuable preliminary insights, we recognise its limitations in capturing the diversity of educational contexts and practitioner perspectives. Future research will address this by recruiting a larger and more diverse participant pool, including educators from varied cultural and institutional backgrounds. Expanding the sample size will provide a stronger statistical foundation and enhance the generalisability of findings. The current study, with its small sample size and single geographic focus (New Zealand), provides valuable insights but also highlights the need for further testing across more diverse contexts. For instance, the practitioners involved in this evaluation offered valuable feedback, particularly regarding the role of AI in administrative tasks. However, as the practitioners themselves represented a relatively homogenous group in terms of geography and institutional background, future iterations of this study must expand the sample to include educators from different cultural and institutional settings. This expansion will provide a more robust foundation for validating the AITTTS and ensuring its generalisability across a wider range of educational contexts.

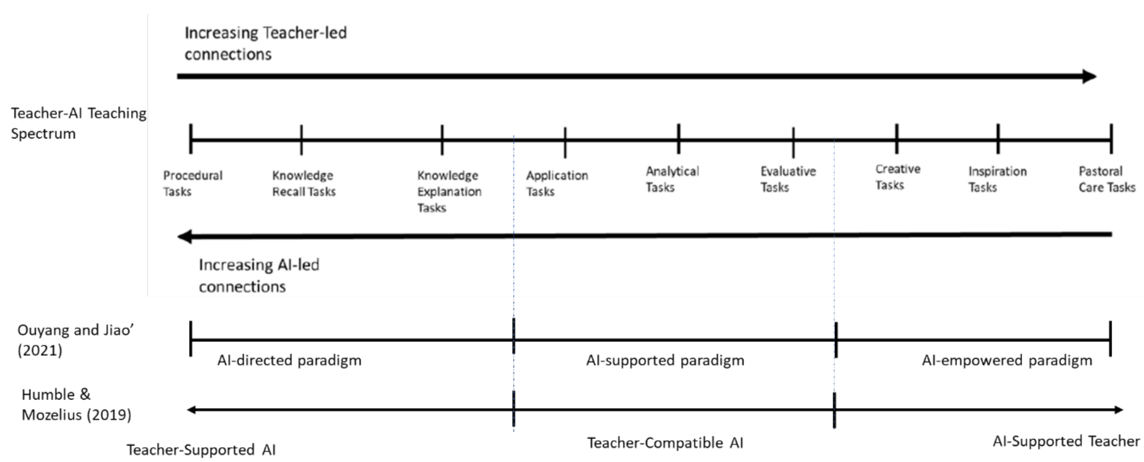
Moreover, future studies should incorporate a longitudinal approach to capture how the AITTTS adapts to new AI developments and educational needs over time. As AI technologies continue to evolve, so too must the framework. For example, while current AI systems may struggle with creativity and emotional intelligence, future advancements could shift the boundaries of what AI is capable of in these areas. A longitudinal study would allow researchers to track these changes and refine the framework accordingly.

Situating the AI-Teacher Teaching Spectrum within the broader literature

Phase 1 - The systematic literature review phase of this study examined fourteen relevant studies to understand the role of AI within the AI-teacher relationship. However, only three of these studies—Lodge, Thompson & Corrin (2023), Ouyang & Jiao (2021), and Humble & Mozelius (2022)—were discussed in detail. These three were selected because they offered the most direct insights into task delegation between AI and teachers, specifically aligning with the structure of the AITTTS. The remaining ten studies, while informative, focused more on broader AI applications or tangential topics that did not directly contribute to the primary goal of validating the framework. To maintain coherence and focus, these studies were summarised in the background section rather than integrated into the findings. This approach ensures that the discussion remains tightly focused on evaluating the framework's relevance to specific AI-teacher interactions.

Figure 2

Situating the AI-Teacher Teaching Tasks Spectrum



The alignment of the AITTTS with existing literature is crucial for situating the framework within broader educational theories. The systematic literature review supporting this study focused on narrowing an initial set of 86 records to three key studies that explicitly addressed AI-teacher task delegation. This selection was guided by the study's pilot nature, which prioritised direct relevance to the AITTTS framework over broader contextual insights. While this focused approach allowed for detailed engagement with pertinent studies, it inherently limits the generalisability of the findings. Future research should address this by broadening the scope of the literature review to incorporate additional studies, particularly those examining emerging applications of AI in education. The categorisation of knowledge recall and explanation tasks as AI-led also fits neatly within existing paradigms (Figure 2). However, more complex tasks—those requiring creative input or pastoral care—are still largely the domain of human teachers. AI's limitations in handling these tasks are well-documented (Humble & Mozelius, 2022), and the AITTTS's distinction between AI-assisted and AI-led tasks reflects this understanding.

Nonetheless, situating the AITTTS "perfectly" within the broader literature oversimplifies the complex interplay between AI and human teachers. The literature indicates that while AI can handle many administrative and procedural duties, its role in more nuanced tasks like creativity and emotional support remains limited. Therefore, a more critical engagement with these complexities is necessary to fully articulate how the AITTTS both aligns with and diverges from existing frameworks. Future research should explore these areas further, particularly to better understand the boundaries of AI's involvement in educational tasks that demand human intuition, creativity, and emotional intelligence.

As such, it can be clearly seen that the framework not only does not contradict the established frameworks, it sits well with the broader literature and is validated and supported by them.

Methodological Reflection

The methodology employed in this study provided a useful starting point for evaluating the AITTTS framework. However, it is important to acknowledge several limitations that affect the study's generalisability and robustness. The small sample size of three academic practitioners, while sufficient for an initial exploration, limits the depth and breadth of the insights that can be drawn. Although Hennink and Kaiser (2022) suggest that 3-5 experts can be appropriate for early-stage validation, a larger, more diverse sample would offer a stronger foundation for drawing conclusions about the AITTTS's applicability across different educational contexts.

Additionally, the geographic and institutional homogeneity of the participants presents a limitation. All participants were based in New Zealand, which may reflect localised educational practices and perspectives that are not necessarily transferable to other regions. Future research should include practitioners from diverse geographic, cultural, and institutional backgrounds to ensure that the framework can be applied globally. This would provide a more comprehensive evaluation of how the AITTTS interacts with varying educational environments and AI integration strategies.

The reliance on Fleiss' Kappa to measure agreement among the participants was also a key aspect of the methodology. While this metric provided a quantitative measure of consensus, it did not capture the more nuanced disagreements among the practitioners, particularly regarding AI's role in creative and pastoral tasks. In hindsight, incorporating qualitative methods, such as in-depth interviews or focus groups, could have enriched the data by providing deeper insights into the reasons behind the differing opinions. A mixed-methods approach would allow for a more comprehensive understanding of both the areas of agreement and divergence, thus providing a richer foundation for refining the AITTTS framework.

Furthermore, the survey instrument itself, although designed to capture practitioners' evaluations of the AITTTS, could be expanded in future iterations. The use of a semi-structured survey was intended to gather both quantitative and qualitative data; however, in practice, the qualitative component was underrepresented. Moving forward, enhancing the survey design to include more open-ended questions could facilitate a more detailed exploration of practitioners' experiences and perspectives, particularly regarding AI's role in tasks like creativity, critical thinking, and pastoral care.

Finally, it is important to reflect on the generalisability of the study's findings. While the results offer valuable insights into how practitioners perceive the AITTTS, the small sample size and limited geographic context constrain the extent to which these findings can be generalised to other educational settings. Future research should address these limitations by expanding the sample size, increasing the diversity of participants, and employing more robust mixed methods approaches to gain a fuller understanding of the framework's applicability across different contexts.

Interpretation of the Survey Results

To address RQ2, practitioner feedback was analysed using Fleiss' Kappa, resulting in a substantial agreement score of 0.73 (Landis & Koch, 1977) (Table 6). While the small sample size of three practitioners limits the generalisability of these results, this score suggests that the AITTTS framework shows promise in defining the roles of AI and human teachers in educational tasks.

The decision to use Fleiss' Kappa was made to provide a structured, statistical measure of agreement among the practitioners. Given the small sample size, this metric offers a straightforward way to quantify consensus, making it useful for an initial, exploratory validation of the framework. However, Fleiss' Kappa does not fully capture the differences in practitioner perspectives. For example, one practitioner strongly believed AI could foster creativity, a view not shared by the others, indicating divergent opinions on creative tasks. Similar differences emerged regarding inspirational tasks, where AI's role was debated among the participants. [insert quotes]

Future research should combine quantitative and qualitative methods to explore these divergences more deeply at scale. This approach will allow for a fuller understanding of AI's limitations in higher-order tasks like creativity and emotional intelligence, providing a more balanced assessment of the AITTTS framework's applicability in various educational settings.

Table 6

Fleiss' κ values and strength of agreement

Fleiss' κ	Strength of Agreement
0.00	Poor
0.1–0.20	Slight
0.21–0.40	Fair

0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

Implications of the AI-Teacher Teaching Task Spectrum

The substantial agreement among practitioners regarding the validity of the AITTTS highlights its potential as a guiding framework in the evolving landscape of AI in education. However, the variation in responses, particularly concerning AI's role in creative and inspirational tasks, suggests that the integration of AI in these areas remains contentious. This underscores the need for a more nuanced understanding of AI's capabilities and limitations, particularly in tasks that require emotional intelligence and creativity—domains where human teachers continue to excel.

Policy

The substantial consensus among practitioners regarding the AITTTS highlights its potential as a foundational tool for policymakers who are grappling with how best to integrate AI into education. As AI technologies continue to evolve and permeate educational systems, there is a pressing need to establish clear guidelines and standards for their use. The AITTTS offers a framework that could help shape national and regional education policies by specifying which tasks are best suited to AI and which should remain the domain of human educators.

One immediate policy application could be the development of AI task delegation guidelines. National education boards could adopt the AITTTS to determine the extent to which AI can be embedded in the classroom without undermining the essential role of teachers. For example, policies could be created to delegate repetitive or procedural tasks, such as grading and scheduling, to AI systems. These tasks are time-consuming and often detract from teachers' core responsibilities of fostering student engagement and creativity. By freeing teachers from these administrative burdens, policy could encourage a shift towards more human-centred, creative, and emotionally intelligent teaching.

Furthermore, the AITTTS can serve as a basis for AI ethics frameworks in education. As AI becomes more prevalent, concerns about its ethical implications, such as student privacy, algorithmic bias, and over-reliance on technology, have grown. The AITTTS clearly delineates the boundaries of AI's involvement, advocating for its use in supporting roles rather than replacing the human elements essential to education, such as pastoral care and creativity. Policymakers could use the AITTTS to craft ethical guidelines ensuring AI is used responsibly and does not infringe on the critical relational aspects of teaching. These guidelines could include stipulations on data privacy, transparency in AI-driven decision-making, and limitations on the role of AI in tasks that require human empathy and emotional intelligence.

At a broader level, education ministries could integrate the AITTTS into curriculum reform initiatives. By providing a structured approach to AI integration, the framework could help educational leaders envision a future where AI enhances—rather than diminishes—teacher-student interaction. For instance, AI could handle the administrative tasks involved in managing large classes, enabling teachers to focus on personalised learning strategies that AI systems are not capable of delivering. Policymakers could also incorporate the AITTTS into teacher training programmes, ensuring that future educators are prepared to collaborate effectively with AI and understand its role in the modern classroom.

Practice

The practical implications of the AITTTS are equally significant, offering educators and school administrators a structured, evidence-based approach to AI task delegation. In day-to-day teaching, the AITTTS could be employed to streamline classroom management, particularly by automating routine administrative tasks. Schools could implement AI systems for grading, attendance tracking, and lesson scheduling, which would reduce the administrative load on teachers and allow them to concentrate on the aspects of teaching that require creativity, problem-solving, and personal interaction. This division of labour would not only improve teacher efficiency but also enhance the quality of education by enabling more tailored and engaging student interactions.

However, for the AITTTS to be successfully implemented in practice, it is essential that schools invest in comprehensive teacher training. Teachers need to develop the digital literacy required to work alongside AI systems, and they must be trained to understand AI's capabilities and limitations. This training should focus on how teachers can effectively integrate AI into their workflows without ceding control of key educational tasks. For instance, teachers should learn how to use AI to assist with procedural tasks like marking and data analysis, while retaining their central role in providing feedback, engaging students, and facilitating higher-order thinking.

In addition to teacher training, schools need to develop robust AI systems capable of handling nuanced educational tasks without overstepping ethical boundaries. For example, while AI can assist with grading factual content, it may struggle with more subjective tasks like grading essays or creative projects, where human insight is crucial. The development of AI tools that can assist teachers in these areas without taking over their decision-making processes is an area ripe for innovation. The AITTTS could provide a framework for developing AI technologies that are specifically designed to complement, rather than replace, human educators.

Moreover, practical implementation would benefit from the adoption of the AITTTS as a strategic tool for professional development. Schools could use the framework to assess their current use of AI, identify gaps in their AI integration, and develop strategies for improving teacher-AI collaboration. By using the AITTTS to audit AI's role in the classroom, school leaders can ensure that AI is being utilised in ways that enhance learning outcomes without encroaching on areas where human teachers excel. This could lead to more thoughtful, measured adoption of AI in schools, avoiding the risks of either over-reliance on technology or underutilisation of its capabilities.

In terms of classroom practice, the AITTTS could also promote greater equity in education by enabling teachers to spend more time with students who need additional support. With AI managing routine tasks, teachers could focus their attention on students who require personalised instruction or pastoral care, addressing learning gaps and fostering a more inclusive learning environment. This would be especially beneficial in large classes or schools with limited resources, where teachers are often overburdened and unable to provide individualised attention to every student.

Finally, the AITTTS framework has the potential to influence the design of future AI tools for education. Developers could use the AITTTS to create AI systems that are specifically tailored to handle the procedural and administrative tasks outlined in the framework, leaving the more complex, creative, and relational tasks to human teachers. By focusing AI development on areas where technology can provide the most value, while ensuring that the human element remains central to the teaching process, the AITTTS could shape the next generation of educational technology.

Limitations

Sample Size and Statistical Power

The pilot study design inherently limits the scope and generalisability of the findings. The small sample size and geographic focus were intentional, reflecting the study's aim to explore the feasibility of the AITTTS framework and identify areas for further investigation in future, larger-scale studies. While early-stage validation studies often recommend a minimum of 3-5 experts (Hennink & Kaiser, 2022), this small sample size limits the ability to generalise the findings. Small samples are frequently criticised for their limited statistical power and inability to detect nuanced differences (Vasileiou et al., 2018). Although Fleiss' Kappa was used to measure agreement, providing a structured tool for quantifying consensus, a larger sample would offer a more robust statistical foundation and allow for greater diversity in perspectives (McHugh, 2018).

Demographic and Geographic Constraints

Another limitation stems from the demographic and geographic homogeneity of the participants, all of whom were based in New Zealand. This limits the cultural and institutional diversity of the findings, as educational practices vary significantly across different regions. Localised policies, pedagogical approaches, and institutional structures could influence the applicability of the AITTTS framework. Thus, relying on a sample from a single geographic region makes the conclusions context-bound and restricts their generalisability to other educational systems (Dawadi et al., 2021). Future research should incorporate practitioners from diverse geographic and cultural backgrounds to ensure broader applicability.

Fleiss' Kappa and Agreement Limitations

While Fleiss' Kappa was used to assess inter-rater agreement among the practitioners, the small sample size restricts the robustness of this analysis (McHugh, 2018). The limited number of participants may not capture the full range of opinions, particularly regarding more contentious tasks, such as creativity and emotional intelligence. A mixed-methods approach combining quantitative and qualitative insights, such as open-ended interviews or focus groups, would provide a richer understanding of practitioner perspectives (Weyant, 2022).

Future Directions for Research

To address these limitations, future research should focus on increasing the sample size and including participants from a wider range of educational and cultural backgrounds. This would enhance the generalisability of the findings and provide a more comprehensive view of how the AITTTS framework can be adapted to various educational settings. Incorporating longitudinal approaches would also allow researchers to capture the evolving use of the AITTTS over time and offer insights into its application across diverse educational contexts (Gerring, 2017). Comparative studies between different countries and educational systems would help identify cultural and institutional factors that influence the framework's applicability (Palinkas et al., 2019).

Conclusion

In conclusion, the AI-Teacher Teaching Tasks Spectrum represents a significant step forward in understanding the complex interplay between AI and human educators. By providing a structured framework that aligns AI capabilities with pedagogical needs, the AITTTS offers a pathway for integrating AI into education in a way that enhances rather than diminishes the role of teachers. As AI continues to evolve, frameworks like the AITTTS will be crucial in ensuring that education remains a fundamentally human endeavour, supported by, but not supplanted by, technology.

Acknowledgements

The authors disclose that they have no actual or perceived conflicts of interest. The authors disclose that they have not received any funding for this manuscript beyond resourcing for academic time at their respective university. The authors have produced this manuscript without artificial intelligence support.

References

Alharmoodi, B. Y. R., & Lakulu, M. M. (2022). The Formulation and Validation of a Conceptual Framework for the Transition from E-government to M-government. European Journal of Interdisciplinary Studies, 8(1), 23–34. https://doi.org/10.26417/502cxc20

Ashri, D., & Sahoo, B. P. (2021). Open Book Examination and Higher Education During COVID-19: Case of University of Delhi. Journal of Educational Technology Systems, 50(1), 73–86. https://doi.org/10.1177/0047239521013783

Bradford, N., Chambers, S., Hudson, A., Jauncey-Cooke, J., Penny, R., Windsor, C., & Yates, P. (2019). Evaluation frameworks in health services: An integrative review of use, attributes and elements. Journal of clinical nursing, 28(13-14), 2486–2498. https://doi.org/10.1111/jocn.14842

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 1–12. https://doi.org/10.1080/14703297.2023.2190148

Dawadi, S., Shrestha, S., & Giri, R. A. (2021). Mixed-Methods Research: A Discussion on its Types, Challenges, and Criticisms. Journal of Practical Studies in Education, 2(2), 25-36 DOI: https://doi.org/10.46809/jpse.v2i2.20

De Freitas, J., Uğuralp, A. K., Oğuz‐Uğuralp, Z., & Puntoni, S. (2023). Chatbots and mental health: Insights into the safety of generative AI. Journal of Consumer Psychology. https://doi.org/10.1002/jcpy.1393

Dumas, J., Sorce, J., & Virzi, R. (1995). Expert Reviews: How Many Experts is Enough? Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 39(4), 228-232. https://doi.org/10.1177/154193129503900402

Fernández-Gómez, E., Martín-Salvador, A., Luque-Vara, T., Sánchez-Ojeda, M. A., Navarro-Prado, S., & Enrique-Mirón, C. (2020). Content Validation through Expert Judgement of an Instrument on the Nutritional Knowledge, Beliefs, and Habits of Pregnant Women. Nutrients, 12(4), 1136. https://doi.org/10.3390/nu12041136

Friedman, S., Hurley, T., & Fishman, T. (2020). COVID-19's impact on higher education: Strategies for tackling the financial challenges facing colleges and universities. https://doi.org/https://www2.deloitte.com/content/dam/Deloitte/us/Documents/public-sector/us-gps-covid-university-finance.pdf

Fölstad, A., Araujo, T., Law, E. L., Brandtzæg, P. B., Papadopoulos, S., Reis, L., Báez, M., Laban, G., McAllister, P., Ischen, C., Wald, R., Catania, F., Von Wolff, R. M., Hobert, S., & Luger, E. (2021). Future directions for chatbot research: an interdisciplinary research agenda. Computing, 103(12), 2915–2942. https://doi.org/10.1007/s00607-021-01016-7

Gerring, J. (2016). Case study research. https://doi.org/10.1017/9781316848593

Hasanpoor, E., Hallajzadeh, J., Siraneh, Y., Hasanzadeh, E., & Haghgoshayie, E. (2019). Using the methodology of systematic review of reviews for Evidence-Based Medicine. Ethiopian Journal of Health Sciences, 29(6). https://doi.org/10.4314/ejhs.v29i6.15

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine, 292, 114523. https://doi.org/10.1016/j.socscimed.2021.114523

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and Implications for Teaching and Learning.

Koh, J., Cowling, M., Jha, M., & Sim, K. N. (2022). Collaborating with AIed for better student-teacher reconnection. ASCILITE Publications, e22126. https://doi.org/10.14742/apubs.2022.126

Koh, J., Cowling, M., Jha, M., & Sim, K. N. (2023). The Human Teacher, the AI Teacher and the AIed-Teacher Relationship. Journal of Higher Education Theory and Practice, 23(17). https://doi.org/10.33423/jhetp.v23i17.6543

Leung L. Validity, reliability, and generalizability in qualitative research. J Family Med Prim Care. 2015 Jul-Sep;4(3):324-7. doi: 10.4103/2249-4863.161306. PMID: 26288766; PMCID: PMC4535087.

Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. Australasian Journal of Educational Technology, 39(1), 1–8. https://doi.org/10.14742/ajet.8695

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson.

McHugh M. L. (2012). Interrater reliability: the kappa statistic. Biochemia medica, 22(3), 276–282. https://pubmed.ncbi.nlm.nih.gov/23092060/

Nowell, L., Norris, J. M., White, D., & Moules, N. J. (2017). Thematic analysis. International Journal of Qualitative Methods, 16(1), 160940691773384. https://doi.org/10.1177/1609406917733847

Palinkas, L. A., Mendon, S. J., & Hamilton, A. B. (2019). Innovations in mixed methods evaluations. Annual Review of Public Health, 40(1), 423–442. https://doi.org/10.1146/annurev-publhealth-040218-044215

Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. Computer Applications in Engineering Education, 28(6), 1549–1565. https://doi.org/10.1002/cae.22326

Polit, D. F., & Beck, C. T. (2021). Nursing research: Generating and Assessing Evidence for Nursing Practice. LWW.

Selwyn, N. (2019). Should robots replace teachers?: AI and the Future of Education. Polity.

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner–instructor interaction in online learning. International Journal of Educational Technology in Higher Education, 18(1), 54. https://doi.org/10.1186/s41239-021-00292-9

Shen, L., & Su, A. (2020). The Changing Roles of Teachers With AI (pp. 1–25). https://doi.org/10.4018/978-1-5225-7793-5.ch001

Tiddi, I., & Schlobach, S. (2022). Knowledge graphs as tools for explainable machine learning: A survey. Artificial Intelligence, 302, 103627. https://doi.org/10.1016/j.artint.2021.103627

Tufford, L., & Newman, P. (2010). Bracketing in qualitative research. Qualitative Social Work, 11(1), 80–96. https://doi.org/10.1177/1473325010368316

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Medical Research Methodology, 18(1). https://doi.org/10.1186/s12874-018-0594-7

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. Expert Systems With Applications, 252, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. Learning Media and Technology, 45(2), 107–114. https://doi.org/10.1080/17439884.2020.1761641

Weyant, E. (2022). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 5th edition. Journal of Electronic Resources in Medical Libraries, 19(1–2), 54–55. https://doi.org/10.1080/15424065.2022.2046231

Young, W. (2022). Virtual Pastor: virtualization, AI, and pastoral care. Theology and Science, 20(1), 6–22. https://doi.org/10.1080/14746700.2021.2012915

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? In International Journal of Educational Technology in Higher Education (Vol. 16, Issue 1). Springer Netherlands. https://doi.org/10.1186/s41239-019-0171-0