Giving the Bird Wings: AI as a Mathematical Coach for Year 8 Girls

Susan Jackson

Diocesan School for Girls, Auckland, New Zealand

Mā te huruhuru, ka rere te manu,

With feathers, the bird flies.[1]

Abstract

Many studies highlight the importance of fostering mathematical confidence in early adolescence, particularly among girls, who often experience a decline in self-efficacy despite capable performance (Zander et al., 2020). My action research explored how using AI-generated prompts to guide Year 8 girls through the MPTC (Make sense, Plan an approach, Take action, Convince yourself and others) mathematical investigation cycle built confidence in non-routine mathematics challenges. Over 16 weeks, 22 Year 8 students engaged with a custom AI chatbot designed to scaffold their work on non-routine, challenging problems using the MPTC cycle. Using a mixed-methods approach, the study employed Mertler's (2020) inductive analysis to identify patterns in the collected data. I constructed four key themes: refined AI-generated prompts build student trust and confidence; confidence grows when AI provides scaffolds rather than solutions; students internalise the MPTC cycle through repeated guided use; and non-judgmental AI interaction increases confidence and fosters agency. My findings revealed that a carefully designed AI chatbot significantly increased participants' confidence when the tool prioritised conciseness, developmental appropriateness, and step-by-step metacognitive checking. This research demonstrates that AI, designed as a pedagogical coach rather than an answer generator, can empower girls to embrace productive struggle and see themselves as capable mathematicians. Implications for practice include the critical importance of prompt engineering and the potential for AI to build mathematical confidence in early adolescent girls.

1 Whakataukī are traditional Māori proverbs passed down through generations; in Te Ao Māori, they emphasise that learners do not succeed in isolation but can realise their potential when provided with the right support and tools.

Glossary

Confidence: A learner's willingness to attempt challenging problems without the fear of making mistakes. It is demonstrated through specific behaviours, including voluntarily asking clarifying questions, articulating reasoning using mathematical language, and expressing positive self-reflection regarding one's ability to learn and improve.

Scaffolding: An instructional technique in which a teacher (or, in this case, an AI chatbot) provides successive levels of temporary support to help students reach higher levels of comprehension and skill acquisition that they may not be able to achieve independently. As learners improve, this support is adjusted to allow them to take increasing ownership of the learning process.

AI Chatbot: A custom-built artificial intelligence AI chatbot, created using Playlab (Playlab, n.d.), that acts as a supportive guide during problem-solving.

MPTC Cycle: A four-step mathematical investigation process where students make sense of a problem, plan an approach, take action, and justify their findings. It emphasises connecting concepts, using representations, and explaining reasoning to convince themselves and others. (New Zealand Ministry of Education, n.d.).

Non-Routine Problem-Solving: The process of working through a mathematical challenge that involves choosing strategies, applying them, and checking the result. It often requires persistence, creativity, and logical thinking.

Productive Struggle: A learning approach where students engage with challenging tasks that require effort, persistence, and thinking, supporting deeper understanding as they make sense of ideas, test strategies, and refine their reasoning rather than being given immediate solutions.

Self-Efficacy: A student's belief in their ability to succeed in specific situations or accomplish a task, which influences their motivation and persistence in mathematics.

Giving the Bird Wings: AI as a Mathematical Coach for Year 8 Girls

In the New Zealand education landscape, Year 8 marks a critical juncture for female students. It is a developmental stage during which attitudes towards mathematics often crystallise, and research consistently highlights a confidence gap, with girls underestimating their competence, even when their academic performance is high (Mozahem et al., 2020; Zander et al., 2020). This gap in self-efficacy often manifests as a reluctance to engage with non-routine challenges, the "messy" problems that require persistence, risk-taking, and critical thinking.

The motivation for this action research stemmed from observing capable students in my classroom opt out when faced with unfamiliar mathematical problem-solving and investigations. They possessed the content knowledge but lacked the structural confidence to begin. Simultaneously, the rapid emergence of generative AI presented a unique opportunity. Could this technology, often feared as a tool for cheating, be repurposed as a non-judgmental coach?

In this study, I explored the intersection of structured pedagogy and emerging technology. Specifically, my research question was: “How does using AI-generated prompts to guide students through the MPTC mathematical investigation cycle foster confidence in tackling non-routine mathematics challenges among Year 8 girls?” By programming an AI chatbot to guide students through the MPTC cycle, this project aimed to shift student reliance from teacher-led instruction to autonomous, supported inquiry.

This study follows Mertler's (2020) action research model, a cyclical, practitioner-driven approach that enables teachers to systematically investigate and improve their own practice. Action research was particularly appropriate for this study because it enabled me to respond iteratively to student feedback, refining the AI chatbot in real time based on their lived experiences. Unlike traditional research methods, action research prioritises immediate, meaningful improvements in student learning within authentic classroom contexts, making it well-suited to exploring emerging technologies such as AI, where best practices are still being established. This methodology provided a flexible yet rigorous framework that aligned with my dual goals: enhancing my students' mathematical confidence while developing my own understanding of how AI chatbots can be thoughtfully integrated into mathematics pedagogy.

Literature Review

Developing robust problem-solving skills and confidence is a key challenge in mathematics education. Early secondary school is a crucial stage, especially for girls, who often show a decline in mathematics self-efficacy despite capable performance (Zander et al., 2020). This literature review examines three interrelated topics: (1) structured problem-solving models and metacognition, (2) the mathematics confidence gap among adolescent girls, and (3) the integration of artificial intelligence (AI) as a learning coach. By reviewing recent research in these areas, I established the context for an action research intervention to empower Year 8 girls' confidence in approaching problem-solving and critical-thinking tasks.

Explicit learning models that teach problem-solving processes have long been advocated. Pólya's (1957) four-step method—Understand, Plan, Solve, Review—is a classic example that encourages heuristic reasoning (educated guesses, pattern recognition, analogies) and gives students "a taste of independent thought" (p. v). Similarly, metacognition, initially conceptualised by Flavell (1979) as the monitoring and regulation of cognitive processes, is considered crucial for improving students' mathematical performance and problem-solving abilities (Mohd Saad et al., 2025). Recent research emphasises the critical role of metacognition in fostering persistence during complex problem-solving (Ugpo et al., 2025) and its status as a significant focal point in the field (Hoang et al., 2024). Furthermore, the MPTC cycle operationalises these ideas, guiding students to make sense of a problem, devise a plan, take action, and reflect on the results (New Zealand Ministry of Education, n.d.).

Research evidence strongly supports the explicit teaching of MPTC strategies. William and Maat (2020) found that students without metacognitive guidance often fail to plan, sequence, and monitor their problem-solving effectively, whereas those taught to self-reflect show significant improvement. Hunaifi and Juandi (2023), in a systematic review of studies from

2010 to 2021, further confirmed that sustained use of metacognitive strategies, including heuristic approaches that develop transferable critical-thinking skills, significantly enhances mathematical problem-solving abilities. Together, these findings affirm that embedding structured problem-solving models and metacognitive prompts helps learners become more independent and effective thinkers.

It is also essential to be aware that confidence, as an affective factor, plays a pivotal role in shaping students' learning outcomes. Research consistently shows a gender disparity in mathematical self-belief: girls often rate their abilities lower than boys do, even when their actual performance is similar (Raabe & Block, 2024). This self-efficacy gap can lead capable female students to doubt themselves, shy away from challenges, and ultimately contribute to their underrepresentation in advanced STEM courses (Zakariya, 2022; Zander et al., 2020).

Fortunately, studies also demonstrate that both confidence and agency are malleable, and that fostering them does not mean leaving students to struggle independently. Hattie (2023) emphasises that agency develops through purposeful scaffolding that gradually transfers responsibility to learners. Boaler's (2016) work on "mathematical mindsets" indicates that cultivating a growth mindset, the belief that ability can improve with effort and that mistakes are learning opportunities, can significantly enhance students' achievement and attitudes in mathematics. In one intervention, Boaler (2024) reported that secondary students who experienced a mindset-focused programme scored ~50% higher on a standardised test, on average, with "dramatic shifts" in their confidence and enjoyment of mathematics. Similarly, Martynovsky (2021) observed that middle-school girls who participated in intensive problem-solving enrichment (e.g., Mathematics Olympiad training) developed significantly greater mathematical confidence and improved problem-solving skills. These findings suggest that with supportive, growth-oriented teaching strategies, girls' confidence in mathematics can be substantially improved.

As I considered mathematical problem-solving and the importance of fostering girls' confidence in the subject, the emergence of AI as a learning coach introduced both promising opportunities and necessary cautions. On the positive side, AI tutors and AI chatbots can provide personalised, on-demand support. They adapt to each learner's needs and offer instant feedback or hints, which have been shown to increase student engagement (Vieriu & Petrea, 2025) and reduce mathematics anxiety (Polydoros et al., 2025). Students often feel safer making mistakes with an AI, as a virtual tutor does not judge. One study found that an AI-based tutor that provided private, step-by-step guidance significantly reduced students' anxiety and improved their resilience in mathematics (Polydoros et al., 2025).

On the negative side, researchers caution that over-reliance on AI may undermine the development of critical thinking. If students use an AI tool merely to get answers, they might bypass the productive struggle and reflection that lead to deep understanding. Melisa et al. (2025) warn that while AI can aid learning (for example, by analysing arguments or providing examples), it can also "hinder students' motivation for self-reflection and critical evaluation" (p. 1), if not used judiciously. Moreover, AI solutions are not infallible; they sometimes produce incorrect or biased outputs, so learners must be taught to question and verify the information they provide (Tripathi et al., 2025).

The emerging consensus is that AI works best as a guided aid rather than a replacement for human instruction. Educators recommend framing AI tools as "learning partners" while maintaining the teacher's central role (Awang et al., 2025; Creely & Carabott, 2025). In practice, this means encouraging students to use AI for hints, suggestions, and practice, while teachers guide them in explaining their reasoning and double-checking AI suggestions. With such an approach, AI can enhance personalisation and engagement without displacing the student's active role in problem-solving.

Recent literature further refined my understanding of how AI can scaffold learning, particularly in mathematics. Wang et al. (2025) found that artificial intelligence-based interactive scaffolding (AIIS) significantly enhances goal-setting, self-evaluation, and motivation in secondary students. While their study focused on language, the principles of contingency (adaptive support) and fading (gradual reduction of assistance) are highly transferable to the MPTC cycle in mathematics. Furthermore, Malik et al. (2025) show that large language models can help create effective scaffolds for middle school mathematics when provided with expert-informed prompts and curriculum context, producing tasks rated higher than those created by teachers for alignment and accessibility. This reinforces the idea that the quality of the prompt strongly influences the effectiveness of AI support. Additionally, Lee et al. (2025) highlight that while generative AI can support non-routine thinking, it must be balanced to prevent passivity, a key consideration in my AI chatbot design, which forces students to "take action" before receiving hints towards complete solutions.

The three domains of heuristic learning models, girls' mathematics confidence, and AI coaching, converge in my study. The literature affirms the importance of each area in isolation; however, there is a clear gap in research that integrates them. Few studies have explored how AI-driven scaffolding of metacognitive problem-solving specifically affects girls' confidence and engagement in mathematics. A recent case study of middle school girls using generative AI for mathematical problem-posing found that AI can meaningfully engage learners in metacognitive activities, highlighting opportunities to support confidence and creativity in mathematics (Walkington et al., 2025). By synthesising proven strategies (explicit problem-solving instruction and growth mindset support) with a new AI chatbot, my study aimed to develop resilient, confident problem-solvers. This intervention was directly informed by the literature and addressed both cognitive and affective dimensions by using structured prompts to strengthen problem-solving skills and non-judgmental support to build confidence.

My research aimed to provide practical insights into the balanced use of AI, enhancing learning while fostering the critical thinking and self-belief that students need in and beyond the classroom.

Research Context

The participants in this study were 22 Year 8 girls (aged 12–13 years old) at Diocesan School for Girls in Auckland, New Zealand. The school is an independent, high-performing girls' school with a strong emphasis on empowering young women. It also values innovation in teaching and learning, with a focus on thoughtfully integrating digital technologies to enhance student engagement and learning outcomes.

This class was a mixed-ability group, selected for their diverse levels of confidence in mathematics. Some students readily engaged in extension tasks, while others displayed visible anxiety when confronted with non-routine problem-solving challenges. As their mathematics teacher for the whole year, I had developed a strong understanding of their learning dispositions. I observed their varied responses to mathematical challenges. My teaching philosophy centres on empowering girls to see themselves as capable mathematicians, and I believe that thoughtfully integrated technology can provide effective scaffolding to support both competence and confidence in learning mathematics.

Ethical approval was obtained from the school, and informed consent was obtained from all participants and their parents or guardians before data collection. Students were assured of confidentiality, with pseudonyms used in all reporting. Participants were also informed of their right to withdraw from the study at any time without consequence. All data were stored securely and were accessible only to me.

The Action

My action research took place over 16 weeks during Terms 3 and 4 (July to November 2025). The students had access to devices and were accustomed to digital learning, but before this study, their use of AI in mathematics was negligible or strictly prohibited. The core intervention involved integrating a custom-built AI chatbot, called "MPTC Math Problem Helper," designed to act as a Socratic coach rather than an answer engine. The intervention was designed to align with a mathematical investigation cycle described as MPTC (New Zealand Ministry of Education, n.d).

Unlike standard ChatGPT, which often provides immediate solutions, I programmed the AI chatbot with specific system instructions to guide students through a structured problem-solving process. The workflow began by asking students to make sense of the task by identifying the type of mathematics involved, such as geometry or number, and confirming their understanding of the question text. Next, the AI chatbot encouraged students to plan an approach by suggesting a strategy, for example, “guess and check” or “draw a diagram,” before moving forward. Only after the student attempted a solution would the AI take action, offering a hint or checking their work. Finally, the AI chatbot prompted students to convince others by justifying their answer, reinforcing reasoning and reflection.

The action proceeded in three phases. In the introductory phase (Weeks One and Two), I introduced the MPTC cycle and the concept of AI as a "coach, not an expert." We discussed the potential benefits and pitfalls, explicitly addressing the risk of "hallucinations" or incorrect AI outputs.

During the primary intervention (Weeks Three to Seven), students utilised the AI chatbot during the “Otago Problem Challenge” and “Do Now” sessions (University of Otago, n.d.). Crucially, the action required significant adjustment. Early feedback from student journals indicated that the initial version of the AI chatbot was "too complicated," gave "rows and rows of algebra," and was "unhelpful and annoying." Based on their feedback, I refined the system prompts to enforce a 150-word limit on AI responses, adopt a tone suitable for children, particularly young girls, and ensure the AI asked for confirmations (e.g., "Do you understand this bit?") before moving on. This second version of the AI chatbot became the primary tool for the remainder of the study.

Data Collection

My study used a mixed-methods approach to explore how an AI chatbot influenced Year 8 girls' confidence in mathematical problem-solving. Following Mertler's (2020) action research guidelines, I prioritised qualitative methods to capture student voice and used quantitative tools to track changes over time. All instruments were selected to directly address my research question. To ensure trustworthiness, I applied Mertler's suggested strategy of polyangulation, gathering data from questionnaires, journals, observations, exit tickets, interviews, and work samples to cross-validate findings.

Before introducing the AI chatbot, all 22 students completed a baseline questionnaire combining Likert-scale and open-ended items (e.g., "I feel confident solving challenging maths problems"). Students also attempted non-AI problem-solving tasks and reflected on their strategies, giving insight into their initial confidence levels. To complement this, I conducted short, semi-structured interviews with students at varying levels of confidence.

During the problem-solving activities, which began each lesson, I used a checklist to observe and record student behaviours such as volunteering answers, asking questions, and persisting through mistakes.

My students kept weekly journals, reflecting on their use of the AI coach and their emotional responses to problem-solving. They created their own pseudonyms, which helped students feel a sense of ownership from the outset. Journal prompts encouraged honest feedback, capturing moments of both frustration and breakthrough. To monitor evolving perceptions, I used feedback surveys and exit tickets that asked students to rate the AI's helpfulness and describe whether it increased their confidence. For instance, one prompt asked, "Do you feel you needed less help from the teacher today?" to assess changes in independence levels.

From Week Eight onwards, students periodically completed tasks without AI support to assess independent use of the MPTC cycle, while data collection continued using the methods described above. In Week 16, students completed a post-intervention questionnaire mirroring the baseline survey, including direct-comparison items such as "Using the AI coach made me more likely to try challenging problems." To deepen understanding, I conducted follow-up interviews with a range of students to explore how the use of the AI had influenced their confidence and problem-solving strategies. Notably, students conducted their own video interviews, allowing them to be both comfortable and honest in their feedback. All interviews were audio-recorded and transcribed for accurate analysis.

Data Analysis

I analysed my data using Mertler's (2020) inductive "organise, describe, interpret" technique (pp. 305–311). I coded all qualitative data from my students' journals, interviews, exit tickets and observations to identify recurring themes related to AI use, the MPTC cycle, and confidence. This involved an initial coding of raw data (e.g., student quotes such as "I felt stuck" or "it made it harder") followed by grouping these codes into broader categories, which were then interpreted to develop core themes. For example, initial codes such as "AI too long," "confusing explanations," and "didn't understand" were grouped under the "Barriers to AI Use" category, which contributed to the final theme "Refined AI-Generated Prompts Build Student Trust and Confidence." Similarly, codes such as "I tried again," "I knew what to do next," and "I didn't need help" clustered under "Indicators of Independence," forming the basis for the theme "Confidence Grows When AI Provides Scaffolds, not Solutions."

I analysed quantitative data from surveys, interactions with the AI chatbot, and checklists to identify trends and validate my qualitative findings. For instance, an increase in Likert confidence scores was directly cross-referenced with journal entries describing the positive influence of AI prompts. I kept all data, both raw and organised, in a Google Drive to create a clear audit trail. To ensure the trustworthiness of my interpretations and minimise bias, I collected my own observations and ideas throughout the process. I had colleagues serve as critical friends to review my thematic interpretations, provide a second opinion, and enhance the validity of my findings. To further strengthen credibility, I employed member checking by sharing preliminary findings with students after 16 weeks, allowing them to confirm or challenge my interpretations of their experiences.

Discussion of Results

The analysis of surveys, interview transcripts, and student journals revealed a significant shift in student confidence throughout the intervention. Following data analysis, I constructed four themes that directly connect my findings to my research question. They were: the evolution of trust and confidence through prompt refinement; the role of scaffolding in building confidence; the internalisation of the MPTC cycle; and the shift from anxiety to agency.

Refined AI-Generated Prompts Build Girls' Trust and Confidence

A key finding was that the mere presence of AI did not guarantee confidence; instead, the quality of the interaction was paramount. In the early stages (Weeks Three and Four), student journal reflections conveyed frustration. Student 1 noted, "I did not like using the AI because it made the questions more complicated ... I ended up getting help from my friend." Student 15 described the AI chatbot as "unhelpful and annoying today ... It kept getting everything wrong." These responses align with recent literature warning that inaccurate AI outputs and over-reliance on AI can undermine motivation and critical thinking (Melisa et al., 2025).

Following the introduction of the version two update, which enforced simpler language and step-by-step MPTC cycle checking, student perceptions shifted markedly. In Interview 8, one student commented, “The improvement was huge … it was short, but really detailed … unlike the other one, which was really long and confusing.” Similarly, Student 22 noted, “I could tell that the maths helper had been fixed … it showed me the best way to solve the problem.” Polyangulating these qualitative insights, my classroom observation checklists captured a clear shift in student affect: early hesitation gave way to sustained engagement once prompts were refined, with my notes recording the class as “super happy” and “nearly all students actively on task.” Several students indicated they would have abandoned the AI chatbot had these improvements not occurred. This finding aligns with Malik et al. (2025), who emphasise that Large Language Models require expert-informed prompting and strong curriculum alignment to generate effective scaffolding.

Nevertheless, even with refined prompts, occasional AI malfunctions continued to disrupt student confidence. Student 15, for instance, described ongoing frustration when the AI hallucinated, explaining that even after she corrected it, the AI chatbot would “agree, then get it wrong again.” Such instances highlight how the inherent unpredictability of AI remains a barrier to sustained student trust and reliance.

Girls’ Confidence Grows When AI Provides Scaffolds, Not Solutions

This study examined how AI prompts, rather than solutions, support confidence by fostering perseverance and productive struggle. Confidence gains were closely associated with sustained engagement, with the AI acting as a scaffold that supported persistence without reducing cognitive demand. Students repeatedly described feeling supported while still “doing the work,” a condition linked to the development of mathematical resilience (Boaler, 2016).

Student journals provided explicit evidence of perseverance: Student 2 noted that the AI’s hints “encouraged me to persevere”; Student 7 persisted through four independent attempts before seeking support; and Student 15 reflected that it was the first time she had spent a full 30 minutes on a problem challenge. These behaviours were mirrored in classroom observations, where students who typically disengaged instead persisted through difficulty. Post-intervention survey responses reinforced this shift, with students reporting that they were “less easy to give up” and “don’t give up as fast.” Together, these data suggest that well-designed AI scaffolding can normalise struggle and sustain effort, supporting mathematical perseverance and confidence.

However, not all students experienced this scaffolding as empowering. For some highly independent learners, tightly structured AI support felt restrictive and undermined their sense of cognitive ownership. For example, Student 11 recorded in her journal that the step-by-step MPTC cycle prompts had “too many steps” and she “would have rather done it in [her] head.” This sentiment was echoed by Student 12, who found the steps "unnecessary." These responses highlight an important limitation: while AI scaffolding benefits many, it can feel reductive rather than enhancing to students who already possess strong independent problem-solving instincts.

Repeated Guided Use of the MPTC Cycle Promotes Self-Regulated Learning in Girls

A concern in educational technology is dependency, whether students can perform without the AI chatbot. Early in the project (Interview 2), students admitted they had forgotten what the letters "MPTC" stood for. One student confessed, "I don't think I actually got it stuck in my brain ... because didn't we ask like AI to help us?" However, as the intervention continued, the repetitive structure of the AI prompts helped internalise the cycle. By Interview 7, a student correctly recalled, "Make sense, approach, take action and convince yourself and others," and added, "It's invaluable to extend my learning." Student 22’s journal entry was particularly telling: "I try not to use it all the time because now I know what the cycle is, I can just do that now." This told me that she was aware of *how* she was thinking, not just *what* she was producing. Also, by choosing not to use the AI chatbot "all the time," she was reclaiming her own cognitive heavy lifting.

Overall, this suggests that the AI served as a training-wheel mechanism. The structured prompts modelled the metacognitive process until students adopted it as their own internal monologue. This aligns with Wang et al.'s (2025) findings on the effectiveness of interactive scaffolding in promoting self-regulated learning.

Non-Judgmental AI Interaction Increases Confidence and Fosters Agency in Girls

My student’s emotional responses showed a clear transition from early apprehension to growing confidence throughout the intervention. My initial observation checklist (July 21) noted excitement mixed with hesitation. In contrast, my late-September field notes described the class as “super happy” with nearly all students actively on task, indicating increasing comfort and engagement. This growing confidence was accompanied by a visible increase in perseverance; I observed that students who typically opted out of problem-solving began actively persisting with the questions. Furthermore, the AI fostered meaningful collaboration. Because the AI chatbot often suggested different strategies to different users, it sparked organic peer discussions about which mathematical approach was best. Finally, students demonstrated a new sense of agency over their learning tools, with one student explicitly stating, “I like that we get to make the AI chatbot better for us.” The quantitative findings supported this shift. Pre- and post-intervention questionnaires indicated that most students reported increased confidence in mathematical problem-solving. Most students, who had initially expressed uncertainty about attempting the problems, reported strong confidence in making a start by the end of the project.

Reflections indicated that the AI's non-judgmental nature was a key factor in reducing anxiety. Student 7 wrote, "It didn't use big words, which helped." Another student in Interview 7 mentioned, "I know that, like if I made even like a small spelling mistake, the AI is always there to help." This helped reduce negative feelings about mistakes and eased mathematics anxiety, both of which lower confidence and contribute to fewer girls pursuing STEM (Zander et al., 2020). By lowering the psychological stakes of asking for help, the AI created a psychological safety zone in which non-routine challenges felt accessible. As one student concluded in the final interview, "It made me more excited about problem-solving."

Conclusions

My study demonstrates that using AI-generated prompts to guide Year 8 girls through the MPTC cycle can significantly foster confidence and engagement in non-routine mathematics, provided the AI chatbot is carefully designed. The AI acted not as an oracle of answers, but as a patient, non-judgmental coach that scaffolded the metacognitive process.

My findings highlight several important aspects. First, design matters: generic AI models may overwhelm young students, so incorporating system instructions that enforce brevity and provide step-by-step guidance is essential for maintaining engagement. Second, scaffolding builds independence: rather than creating dependence, the repetitive structure of the AI prompts helped students internalise the MPTC cycle, which, in turn, led to improved independent problem-solving. Finally, the outcomes were significant: the most notable gain was in students' emotional relationship with challenging mathematics. The AI reduced the anxiety associated with "getting started," turning it into a manageable, even productive, part of the learning process.

For educators, this research suggests that AI can be integrated into mathematics classrooms not as a computational shortcut, but as a pedagogical support that reflects effective teaching practices such as scaffolding, questioning, and metacognitive prompting. When thoughtfully designed, AI can support student confidence without diminishing productive struggle. Further research is needed to determine whether these confidence gains persist as students progress into senior secondary mathematics.

While the findings are promising, this study has limitations. Conducted in a single, well-resourced girls' school with high digital literacy, the results may not be transferable to co-educational or under-resourced contexts. The 16-week duration also limits conclusions about long-term confidence retention. Additionally, the focus on non-routine problem-solving means impacts on procedural fluency, conceptual understanding, or high-stakes assessment performance remain unknown. Future research could explore transferability across diverse settings, track students longitudinally, and examine student-designed agentic AI supports beyond AI chatbot-based tools.

Reflection Statement

As a mathematics educator in a girls' school, I embarked on this research with a mix of excitement and apprehension. I had heard too many capable Year 8 girls ask to "opt out" when faced with non-routine problems; students who understood mathematics but lacked the confidence to begin. I was motivated by a desire to help them see themselves as capable mathematicians who can navigate uncertainty rather than avoid it. What I did not anticipate was how much this project would challenge and reshape my own practice.

Navigating the AI frontier proved to be as demanding for me as the mathematical challenges were for my students. Learning to “prompt engineer,” troubleshooting unexpected outputs, and responding to early student feedback, though sometimes blunt, was genuinely humbling. My students helped me redesign the AI chatbot, resulting in shorter responses, a more supportive tone, and a greater emphasis on student agency within the MPTC cycle.

The most rewarding aspect of this work was observing students take increasing control of their learning and grow more independent. Girls who had previously hesitated at the sight of a word problem began to navigate the MPTC cycle with confidence, “chatting” their way through tasks, articulating their thinking, and supporting one another without relying on teacher intervention. Particularly powerful was hearing students who were initially sceptical later reflect on how the revised AI chatbot helped them organise and extend their thinking without doing the work for them. This developing sense of agency, in which scaffolding supported rather than replaced productive struggle, was central to the intervention’s success.

This research reinforced my belief that technology, when grounded in strong pedagogy and strengthened through genuine human connection, can act as a powerful equaliser. Beyond this, it reaffirmed the value of student voice. My students were not passive participants but co-designers, whose insights directly shaped the learning experience. I wholeheartedly thank my remarkable Year 8 students, who gave me the wings to fly. I would also like to express my sincere gratitude to the leadership team at Diocesan School for Girls, and in particular, Mrs Heather McRae, for her unwavering support and for fostering a culture of innovation that made this project possible. I am grateful to my Research Advisor, Leanne Horwitz, for her thoughtful guidance and constructive feedback, and to my GARC colleagues for their collegiality and good humour. Thank you also to my critical friends for their honest and considered perspective.

Ultimately, this project echoes Hattie's (2012) assertion that the greatest impact occurs when teachers become learners of their own teaching and students become teachers of themselves. Through this research, we embodied both roles.

References

Awang, L. A., Yusop, F. D., & Danaee, M. (2025). Current practices and future direction of artificial intelligence in mathematics education: A systematic review. *International Electronic Journal of Mathematics Education, 20*(2), Article em0823. https://doi.org/10.29333/iejme/16006

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. Jossey-Bass.

Boaler, J. (2024). Transforming mathematics education through mindset-based teaching. *Scientia*. https://doi.org/10.33548/SCIENTIA1096

Creely, E., & Carabott, K. (2025). Teaching and learning with AI: An integrated AI-oriented pedagogical model. *The Australian Educational Researcher, 52*, 4633–4654. https://doi.org/10.1007/s13384-025-00913-6

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning* (p. 18). Routledge. https://doi.org/10.4324/9780203181522

Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge. https://doi.org/10.4324/9781003380542

Hoang, T. N., Vu, T. B., & Nguyen, T. T. (2024). Metacognition in mathematics education: From academic chronicle to future research scenario—A bibliometric analysis with the Scopus database. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(4), Article em2427. https://doi.org/10.29333/ejmste/14381

Hunaifi, A., & Juandi, D. (2023). *A systematic review of metacognitive strategies in mathematical problem-solving.* AIP Conference Proceedings, 2734(1), 090041. https://doi.org/10.1063/5.0175086

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery. https://doi.org/10.1145/3706598.3713778

Malik, R., Abdi, D., Wang, R., & Demszky, D. (2025). Scaffolding middle-school mathematics curricula with large language models. *British Journal of Educational Technology.* https://doi.org/10.1111/bjet.13571

Martynovsky, M. (2021). *Effects of Math Olympiad training on middle-school female students' confidence in mathematics* (Master's thesis, Harvard University). Harvard University DASH Repository. https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37367698

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Al Fuad, Z. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal, 14*(1), Article e2025031. https://doi.org/10.22521/edupij.2025.14.31

Mertler, C. A. (2020). *Action research: Improving schools and empowering educators* (6th ed.). SAGE Publications.

Mohd Saad, M. R., Hidayat, R., & Wewe, M. (2025). A meta-analysis of the effect of metacognitive instruction on mathematics achievement. *Cogent Education, 12*(1), Article 2517510. https://doi.org/10.1080/2331186X.2025.2517510

Mozahem, N. A., Boulad, F. M., & Ghanem, C. M. (2021). Secondary school students and self-efficacy in mathematics: Gender and age differences. *International Journal of School & Educational Psychology, 9*(sup1), S142–S152. https://doi.org/10.1080/21683603.2020.1763877

New Zealand Ministry of Education. (n.d.). *The mathematics investigation cycle: Poster and cards.* Tāhurangi. https://newzealandcurriculum.tahurangi.education.govt.nz/the-mathematics-investigation-cycle-poster-and-cards/5637253334.p

Playlab. (n.d.). *Playlab AI platform*. https://www.playlab.ai

Polydoros, G., Galitskaya, V., Pergantis, P., Drigas, A., Antoniou, A.-S., & Beazidou, E. (2025). Innovative AI-driven approaches to mitigate math anxiety and enhance resilience among students with persistently low performance in mathematics. *Psychology International, 7*(2), 46–60. https://doi.org/10.3390/psycholint7020046

Pólya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Doubleday.

Raabe, I. J., & Block, P. (2024). The gendered maths confidence gap, social influence and social integration. *European Societies, 26*(5), 1596–1631. https://doi.org/10.1080/14616696.2024.2349217

Tripathi, T., Sharma, S. R., Singh, V., Bhargava, P., & Raj, C. (2025). Teaching and learning with AI: A qualitative study on K-12 teachers' use and engagement with artificial intelligence. Frontiers in Education, 10, Article 1651217. https://doi.org/10.3389/feduc.2025.1651217

Ugpo, L. A. B., Tubo, M. T., & Yurango, C. P. (2025). Role of metacognitive awareness in enhancing persistence in solving mathematical problems among math majors. *Asian*

Journal of Education and Social Studies, 51(6), 1304–1311. https://doi.org/10.9734/ajess/2025/v51i62075

University of Otago. (n.d.). *Problem challenge introduction*. https://www.maths.otago.ac.nz/pc/

Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences, 15*(3), Article 343. https://doi.org/10.3390/educsci15030343

Walkington, C., Pando, M., Lipsmeyer, L. L., Beauchamp, T., Sager, M., & Milton, S. (2025). Middle school girls using generative AI to engage in mathematical problem-posing. *Mathematical Thinking and Learning*. https://doi.org/10.1080/10986065.2025.2542724

Wang, F., Zhou, X., Li, K., Cheung, A. C. K., & Tian, M. (2025). The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English. *Interactive Learning Environments, 33*(8), 4589-4608. https://doi.org/10.1080/10494820.2025.2470319

William, S. K., & Maat, S. M. (2020). Understanding students' metacognition in mathematics problem solving: A systematic review. *International Journal of Academic Research in Progressive Education and Development, 9*(3), 115–127. https://doi.org/10.6007/IJARPED/v9-i3/7847

Zakariya, Y. F. (2022). Improving students' mathematics self-efficacy: A systematic review of intervention studies. *Frontiers in Psychology*, *13*, Article 986622. https://doi.org/10.3389/fpsyg.2022.986622

Zander, L., Höhne, E., Harms, S., Pfost, M., & Hornsey, M. J. (2020). When grades are high but self-efficacy is low: Unpacking the confidence gap between girls and boys in mathematics. *Frontiers in Psychology, 11,* Article 552355. https://doi.org/10.3389/fpsyg.2020.552355