The Professional Knowledge Required for High-quality AI-generated Mathematics Lesson Planning

Drew Nucci (WestEd), Alex Liu (University of Washington), Min Sun (University of Washington), and Lorraine M. Males (University of Nebraska)

Mathematics teacher educators (MTEs) help teachers learn ambitious mathematics instruction. Of late, many MTEs wonder how generative AI (GenAI) may be useful in this endeavor, but as is true for any tool, users of GenAI must understand their own purposes and a tool's capacities and limitations. We are a team of MTEs and researchers who have partnered with teachers to design GenAI tools for planning high quality lessons. In this article, we discuss the professional expertise required to design and use GenAI tools for planning ambitious mathematics instruction.

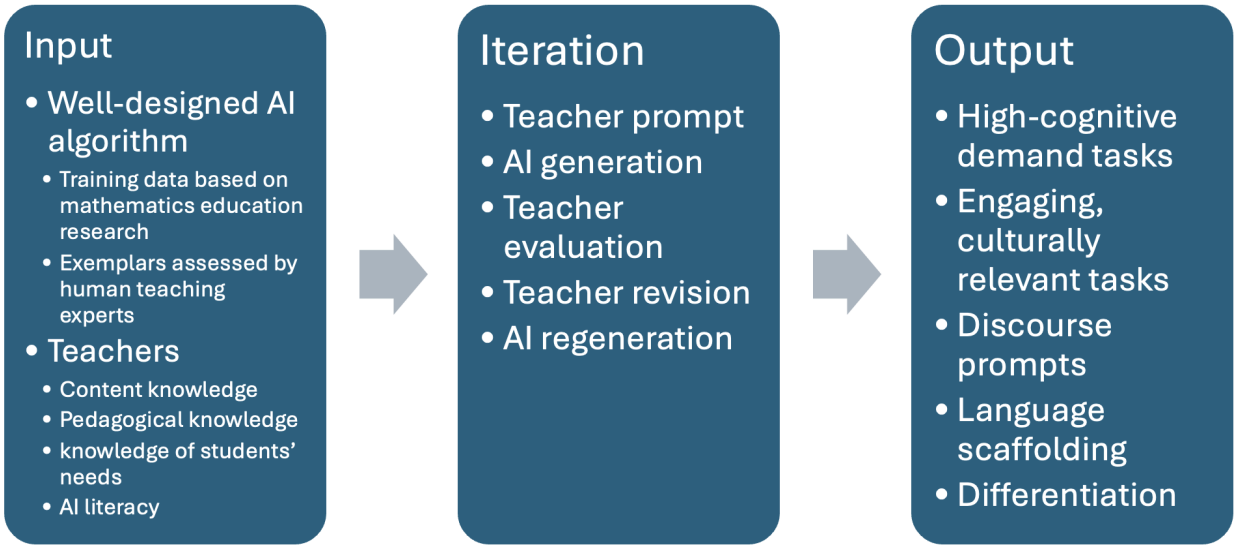
Using GenAI to plan for ambitious mathematics instruction is more involved than typing "Create a rich task about proportionality for 7th graders" into ChatGPT. Ambitious mathematics instruction involves launching level-appropriate, high-cognitive-demand tasks (Stein & Smith, 1998). This requires teachers to understand their students' prior knowledge, believe in their students' mathematical capabilities, and have comprehensive knowledge of mathematical concepts (Wilhelm, 2014). Tasks also must be meaningful to students, which requires teacher knowledge about students. Once an engaging, high-cognitive-demand task is selected, teachers must provide scaffolded launches with language planning for emergent multilingual students (Jackson et al., 2013) and support student task enactment by facilitating student discourse (Kazemi & Hintz, 2014). So, for mathematics instruction to be ambitious, an AI-generated task without the substantial professional knowledge to iterate it is insufficient. Teachers must understand their students, understand how new mathematical concepts connect to students' prior knowledge, understand the language needs of emergent multilingual students, and understand how to facilitate mathematical discourse among students with a variety of learning needs to use GenAI as an effective lesson planning tool. The same professional knowledge must be used to build the standards-aligned GenAI systems upon which teachers can rely. In what follows we describe how GenAI can be created to assist teachers in developing materials for ambitious mathematics teaching and the iterative processes they should employ to maximize GenAI's utility.

GenAI is constrained by the quality and scope of its training data (Zhuo et al., 2023). Without appropriate training GenAI lacks the nuanced understanding that experienced teachers have to create high-quality and inclusive content (Felix, 2020). A GenAI system for ambitious mathematics teaching must be trained using mathematics education research *and* the professional judgment and expertise of experienced teachers. With such training and a teacher's informed iterative use, GenAI can produce materials that support ambitious mathematics teaching.

We propose a model of workflow for mathematics teacher-AI collaborative content design (see Table 1). In this model, input from a well-designed, professionally-informed AI algorithm and teachers' expertise converge through a human-centered iterative process. To develop the AI algorithm, experienced teachers participate in training the GenAI system by inputting example lesson plans that they code with a research-supported rubric. Then teachers who use the GenAI input their professional knowledge: content knowledge, pedagogical knowledge, knowledge of their students' needs, and an understanding of how GenAI systems work. Teachers then engage in iterative rounds of prompting, generation, evaluation, and revision to produce materials that support ambitious mathematics instruction.

Table 1.

Workflow for Mathematics Teacher-AI Collaborative Content Design



Iteration is vital to effective use of GenAI. Iteration begins with a teachers' prompt grounded in criteria for a high-quality task, such as aligning with curriculum standards, promoting conceptual understanding, and facilitating procedural fluency. The initial prompt also needs to provide enough contextual information, such as the intended time duration, grade level, class size, and student demographics to allow the algorithm to adapt content to specific classroom settings. The initial prompt should contain information about the students' prior knowledge for the GenAI to build on, and learning goals and quality criteria, such as promoting critical thinking or encouraging the development of mathematical reasoning, for the GenAI to build toward. If teachers have instructional strategies in mind, they can input an example and then iterate the GenAI tool to produce a similar output.

Once GenAI has produced a lesson plan, teachers should evaluate the output and ask GenAI to improve its output to, for example, provide a plan for more meaningful discourse. GenAI is highly capable of simulating classroom scenarios or generating discourse prompts, which lays the groundwork for developing a student discourse plan

(Dai & Ke, 2022). However, to produce a meaningful discourse plan, the quality, diversity, authenticity, and adaptability of the generated content must be scrutinized. Teachers should look for conceptual questions and prompts to help them assess and advance student thinking. They also should ask GenAI to produce questions that they can pose to groups of students to promote student-student discourse and justification of reasoning.

Knowledgeable teachers can also iterate GenAI for better language scaffolding. For this iterative goal, the design and development of the AI systems have a crucial influence on the generated content. For example, the AI algorithm must be fed training data that adequately represents the students' home languages and cultures to produce useful and relevant output tailored to those students' language needs. The iterative process between teachers and GenAI should go beyond simple tasks like translation to include language scaffolds for task launches and resources for emergent multilingual students to learn and use mathematics vocabulary. Iterating GenAI to create context-rich content that facilitates EMs' access to mathematics concepts and problem solving requires significant pedagogical expertise of teachers.

Teacher expertise is also required to iterate GenAI to plan for students with diverse learning needs. GenAI has been widely studied and used for personalized learning due to its adaptability to different needs (e.g., Reddy et al., 2015), but teachers' knowledge of their students' diverse assets and needs is crucial for effectively leveraging this technology for planning. Equipped with this understanding, teachers can prompt GenAI to generate content tailored to specific student profiles or groups. Teachers then must evaluate the GenAI output for alignment with pedagogical strategies proven effective for students with particular learning needs and iterate, prompting GenAI to improve its outputs to produce inclusive and supportive content for each and every student.

Learning to iterate GenAI to improve its outputs requires some up-front time investment (Mollick, 2024). Once teachers learn to iterate GenAI to build lesson plans for ambitious mathematics teaching, GenAI can be a promising tool for reducing teachers' future lesson planning workload. However, GenAI cannot achieve that promise without teachers' professional expertise. Mathematics teachers' understanding of mathematics, pedagogy, and their students' assets and needs are essential for GenAI to improve lesson planning while also saving teachers' time. We pose these considerations for MTEs because for GenAI to support mathematics teachers in planning for and enacting better tasks and instruction, the essential elements of mathematics teacher education such as developing pedagogical content knowledge, planning for meaningful discourse, and planning for inclusivity are all still very much essential. Without it, teachers will be unable to effectively use GenAI for ambitious mathematics instruction.

The development of the paper was supported by generous funding from the National Science Foundation through Grant #2300291. The opinions expressed are those of the authors and do not represent views of the funder.

References

Ahn, J., Verma, R., Lou, R., Liu, D., Zhang, R., & Yin, W. (2024). Large language models for mathematical reasoning: Progresses and challenges. *arXiv preprint arXiv:2402.00157*.

Dai, C. P., & Ke, F. (2022). Educational applications of artificial intelligence in simulation-based learning: A systematic mapping review. *Computers and Education: Artificial Intelligence*, *3*, 100087.

Felix, C. V. (2020). The role of the teacher and AI in education. *International Perspectives on the Role of Technology in Humanizing Higher Education* (Vol. 33, pp. 33-48). Emerald Publishing Limited.

Jackson, K., Shahan, E. C., Gibbons, L. K., & Cobb, P. (2013). Launching Complex Tasks. *Mathematics Teaching in the Middle School*, *18*(1), 24–29.

Kazemi, E., & Hintz, A. (2014). *Intentional Talk—How to Structure and Lead Productive Mathematical Discussions*. Stenhouse Publishers.

Mollick, E. (2024). *Co-Intelligence: Living and Working with AI*. Random House UK.

Reddy, V. K., Said, L., Sengupta, B., Chetlur, M., Costantino, J. P., Gopinath, A., Flynt, S., Balunaini, P., & Vedula, S. (2015). Personalized learning pathways: Enabling intervention creation and tracking. *IBM Journal of Research and Development*, *59*(6), 4:1-4:14. https://doi.org/10.1147/JRD.2015.2456711

Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School*, *3*(4), 268–275. https://doi.org/10.5951/mtms.3.4.0268

Wilhelm, A. G. (2014). Mathematics teachers' enactment of cognitively demanding tasks: Investigating links to teachers' knowledge and conceptions. *Journal for Research in Mathematics Education*, 45(5), 636-674

Zhuo, T. Y., Huang, Y., Chen, C., & Xing, Z. (2023). Exploring AI ethics of ChatGPT: A diagnostic analysis. https://doi.org/10.48550/arXiv.2301.12867