

## **Mathematics Teachers' Use of Generative AI to Create Active Learning Experiences**

**By:** Candace Walkington<sup>1</sup>, Katie Bainbridge<sup>2</sup>

<sup>1</sup>Southern Methodist University

<sup>2</sup>OpenStax, Rice University

**Keywords:** generative AI; large language models; mathematics education; generative learning

### **Abstract**

Generative AI tools have arisen as important supports to assist teachers with planning instruction. One important way that GenAI can assist teachers in lesson planning is by allowing them to incorporate more “active learning” into their instruction, which we refer to here as “generative learning.” In the present chapter, we explore mathematics teachers ( $n=16$ ) using GenAI to adapt a task they are using in their classroom to incorporate generative learning and also explore GenAI as support as teachers themselves engage in a novel generative learning task related to lesson design. We find that teachers appreciate the creative affordances of generative AI to support generative learning activities but worry about time efficiency, usability, and accuracy, as well as a variety of important big-picture issues. Five implications for policy and practice are given.

### **Introduction**

Generative AI (GenAI) tools have arisen as important supports to assist teachers with planning instruction. Companies like MagicSchool.ai (n.d.), Google Gemini (n.d.), and Khan Academy (n.d.) offer tools that allow teachers to use GenAI to assist them with a wide variety of tasks, from writing items for assessments to brainstorming lesson ideas and formulating lesson plans to adapting instructional materials to their particular learner needs, to understanding common learner misconceptions. However, given the rapid development and scale-up of GenAI for education, little is known about how these tools can be designed and used effectively by teachers. As a result of this lack of knowledge, GenAI use has not become widespread among mathematics teachers. For example, a RAND survey (2024) indicated that 82% of math teachers have *never* used AI tools for mathematics teaching, and only 1% use AI tools “often.” This supports other research showing that mathematics teachers are unsure of how to integrate GenAI into their teaching (Wachira 2025). Teachers may be missing important opportunities to improve their instruction with GenAI.

One important way that GenAI can assist teachers in lesson planning is by allowing them to incorporate more “active learning” into their instruction, which we refer to as “generative learning.” We define generative learning, according to Fiorella and Mayer (2016), as learning that involves “actively making sense of to-be-learned information by mentally reorganizing and integrating it with one’s prior knowledge, thereby enabling learners to apply what they have learned to new situations” (717). In this paper, we explore mathematics teachers using GenAI to

adapt a classroom task to incorporate generative learning. We also explore GenAI as a support as teachers themselves engage in a novel generative learning task related to lesson design.

The concept of generative learning involves facilitating students using learning strategies like (1) summarizing main ideas from learning materials in their own words, (2) making concept maps, knowledge maps, or graphic organizers to convert information to meaningful spatial arrangements, (3) creating drawings or imagined mental images to depict the content of a lesson, (4) answering practice questions about a lesson's material they have learned, (5) engaging in self-explanation of content from learning materials by giving an explanatory account of the material in their own words, (6) teaching the information to others, and (7) engaging in physical motions like object manipulations and gestures during learning, and (8) generating predictions or questions, often prior to learning the target information (Brod 2021; Fiorella and Mayer 2016). GenAI can not only help teachers create activities that incorporate these eight types of learning behaviors but can also provide support as students and teachers engage in these sometimes difficult generative learning practices.

Meta-analyses of generative learning interventions have shown consistent benefits for student outcomes, with a median effect size of  $d = 0.71$  (Mayer 2021). Despite this strong evidence, generative learning activities can be difficult for teachers to incorporate into classroom practice. The creation and implementation of engaging, active learning experiences can be time-consuming for all teachers, especially for those new to the profession (Michael et al., 2023). These activities often require extensive preparation, including creating materials such as flashcards, games, or graphic organizers, as well as anticipating student questions and problems (Fiorella and Mayer, 2020). It may take the teacher conducting a few classes using a generative learning activity for a new activity to run smoothly. Creating new activities or customizing activities for different classes requires an amount of time and creative energy that teachers are reticent to give. This gap between the proven benefits of active learning and its practical application in classrooms provides the perfect opportunity for technology-driven solutions. GenAI could be a useful driver of innovative tools for teachers seeking to quickly and confidently implement a high-quality generative learning activity. For example, AI could generate the appropriate difficulty level of examples used for a flashcard game and execute the tedious formatting work involved in spacing the materials correctly on the page. AI could suggest a list of potential questions students might have about the activity or concept to help teachers prepare. AI could rapidly create more or less scaffolded versions of an existing activity so the teacher can easily match the instructions to a student's ability level. AI could suggest a large number of potential activities well-suited to the subject simply to inspire the teacher's imagination. There are a wide variety of ways in which AI could lower the burden of creating and implementing generative learning activities in the classroom for teachers and thus increase the prevalence of this evidence-based learning strategy.

However, while GenAI has enormous strengths in designing an infinite array of learning materials that are very specific, fine-grained, and targeted, GenAI also has enormous limitations. GenAI does not have the deep understanding that teachers do of the students as people with history and needs (Gómez Marchant and Hardison 2024). It can hallucinate and provide factually incorrect information in lesson materials. It may respond to requests to design fine-grained materials with outputs that are imprecise or wordy, using advanced vocabulary and language

structures (Yen and Hsu 2023), or give ideas that are developmentally inappropriate for students (Sawyer and Aga 2024). GenAI has also been shown to exhibit bias against marginalized groups and a preference for privileged groups in the materials that it outputs (Bender et al. 2021), often related to the biased nature of the training set that the GenAI uses to build its response.

Research on mathematics teachers using GenAI to support their lesson planning has found a variety of utilization schemes. Math teachers can use GenAI as a tutor to gain mathematical or pedagogical knowledge or as a guide to consult for ideas or suggestions. They can use GenAI to simulate fictional students or to organize the task of lesson planning. When making lesson plans with ChatGPT, teachers always adjust the original output through additional prompting to make it suitable for their students' needs, the mathematical content, and their own preferences (Broutin 2024). Mathematics teachers also show adeptness at critiquing the lesson plans generated by GenAI (Aga, Sawyer, and Wolfe, 2024). They identified that lesson plans could be vague and lack details and links to relevant student-facing materials and that even with this generality or vagueness, lesson plans were still often not a fit for their needs and teaching style. When using GenAI to pose mathematics problems specifically, research suggests that teachers use GenAI to adjust difficulty, integrate real-world contexts, and identify potential student difficulties (Biton and Segal 2025). However, teachers may struggle to find and correct errors in mathematics problems they pose with GenAI (Kim, Park, and Joung 2024). Research also suggests that using GenAI to design instructional materials can be time-consuming for teachers due to the amount of iteration needed (e.g., Kuusemets et al., 2024).

We expand this work by looking at using GenAI for teacher lesson planning relating to *generative learning* activities specifically, given the unique challenges of using these approaches in the classroom. Our research questions are:

- (1) What are mathematics teachers' reflections after using GenAI to modify their lesson materials to incorporate more generative learning?
- (2) How do mathematics teachers respond to GenAI as a brainstorming partner as they engage in their own generative learning task?


## **Methodology**

The first group of participants in this study included 11 teachers (9 female, 1 male, 1 other identification) enrolled in an Algebraic Reasoning class at a university in the United States. Eight teachers identified as Asian, two as Hispanic, and one as White. Three teachers were in a pre-service secondary mathematics teacher training program and did not yet have classrooms of their own, while the others typically taught mathematics in levels ranging from 6<sup>th</sup> grade to pre-calculus. The second group of participants in this study included an additional 5 teachers (2 female, 3 male) who were currently teaching Algebra 1 in urban schools, and they were recruited because they were using a particular online curriculum. Both portions of this study were IRB approved.

The 11 teachers enrolled in the course engaged in an activity where they selected a current mathematical task they use in their classroom (the pre-service teachers selected a task they might use or that was used in the class they were observing) and asked ChatGPT to modify it to incorporate active or generative learning (see Figure 1). The teachers were then asked what they

noticed and wondered about ChatGPT’s output and what feedback they had for designers looking to design AI tools to incorporate active learning into instruction. Their written responses and ChatGPT logs (when available) were collected. The 5 additional teachers were engaged in a video-recorded interview where they were asked to use ChatGPT to create an active learning task for some topic they were teaching in their mathematics classroom. They were also asked what they noticed and wondered about ChatGPT’s output and what feedback they had for designers looking to design AI tools to incorporate active learning into instruction. They were further asked about the challenges they saw with using this kind of tool. The interviews were transcribed, and ChatGPT logs (when available) were collected.

Can you modify the following question for active learners? Draw a graph!  
 $y = \frac{1}{2}x - 5$

 To engage active learners, you can transform the task into an exploratory, hands-on activity that promotes discovery and collaboration. Here's a modified version of the question:

---

**Graph Hunt: Bring the Equation to Life!**

**Challenge:** You're an explorer searching for the hidden graph of the equation  $y = \frac{1}{2}x - 5$ . Follow these steps to reveal the graph:

1. **Table Detective:** Create a table of values. Pick at least 5 different values for  $x$  (including negatives!). Calculate the corresponding  $y$ -values.
2. **Point Tracker:** Plot the points you found on a coordinate plane. Use a ruler to connect them and reveal the graph.
3. **Slope Sleuth:** Use the slope  $\frac{1}{2}$  and  $y$ -intercept  $-5$  to double-check your work.

**Active Twist:**

- Work in pairs or small groups to complete the graph. Compare your graphs with your classmates.
- **Bonus Question:** Imagine this graph represents a ski slope. What would happen if the slope was steeper (e.g.,  $y = x - 5$ ) or flatter (e.g.,  $y = \frac{1}{4}x - 5$ )? Sketch these graphs and discuss!

---


**Hands-On Options:**

- Use graphing tools like Desmos or graph paper for physical plotting.
- Turn it into a relay race: Each team member calculates one point or plots one step.
- Let students color-code their graph lines for creativity and better visibility.

Would you like suggestions for group activities or digital tools to further enhance this?

**Figure 1. Example of log of teachers’ exchange with ChatGPT**

As a follow-up, the 11 teachers in the course engaged in an activity where ChatGPT was used as a brainstorming partner when confronting a generative learning task. The teachers were asked to select a picture of their surroundings or local community and come up with mathematics questions that could be posed based on this picture (see Figure 2). The idea was that this picture and the accompanying questions could then be given to the students in their classes to promote mathematical reasoning and connect mathematics to the real world. After teachers posed their questions based on the image, they were asked to give the picture to ChatGPT and have ChatGPT pose mathematical questions based on the image. The teachers were then asked to give written feedback on what they thought of using ChatGPT as a brainstorming partner for this generative learning task, including its pros and cons.

Image	Teacher-Generated Math Question	ChatGPT-Generated Math Question
	<p>If the first tier can hold 100 visitors and each higher tier holds 10 fewer visitors, how many visitors can the building hold?</p>	<p>Geometric Sequences and Series: If each successive tier of the tower decreases in size by a fixed percentage, and the base tier has an area of 10,000 square feet, can you write an expression for the area of the 6th tier? Assume the reduction in area is 20% for each higher tier.</p>

**Figure 2. Generative learning activity (i.e., “generating questions”) completed by teachers in the study**

We used the qualitative method of reflexive thematic analysis to code teacher response data (both interview and written responses; Braun and Clarke 2019). In reflexive thematic analysis, codes capture interesting features of the data, and themes, patterns of shared meaning, are built from individual codes. This method is used to understand themes that occur across a dataset (Braun and Clarke 2012), with the data coming from a variety of sources. Coders iteratively develop the coding categories and an accompanying codebook. Themes arose “at the intersection of the researcher’s theoretical assumptions, their analytic resources and skills, and the data themselves” (Braun and Clarke 2019, 14). The data was coded by the first author, and codes were reviewed by the second author. Some limitations of the methods used for this study are that the sample was a small convenience sample and written feedback was relied on rather than interview data for many of the included teachers.

**Findings**

***RQ1: Teachers’ Reflections on GenAI Output***

When teachers were asked what they noticed and wondered about GenAI’s output when it created generative learning lesson plans (Figure 1), they made wide-ranging observations (Tables 1-2). Many observations were positive (Table 1). Teachers appreciated that the activities proposed by ChatGPT were often hands-on and collaborative: “It is a good way to learn with something touchable, and has a time to work with peers.” The teachers also appreciated that the generative learning activities were engaging and relatable: “I would definitely use them because they’re all relatable, and you know they have the real-world examples tied into it.” The teachers liked that the generative learning activities were open-ended and exploratory, with one teacher describing how “I like how the question was made into more of an exploration question.” They also liked that the activities could be differentiated, with one teacher saying, “I noticed that it [ChatGPT] differentiated the problem, adding more details.” Finally, teachers noticed that ChatGPT made use of multiple representations in mathematical concepts.

**Table 1. Teachers’ reflections on positive elements of using ChatGPT to enhance generative learning**

<b>Theme</b>	<b>Description</b>
1. Hands-on/interactive	In this theme, the teacher expressed appreciation for the activity that ChatGPT created incorporating hands-on learning and allowing for student interactions with content
2. Peer collaboration	In this theme, the teacher expressed that the activity ChatGPT created allowed for students to collaborate with each other in generally meaningful ways, and sometimes even incorporated peer teaching.
3. Open-ended/exploratory	In this theme, the teachers expressed that they appreciated that the tasks generated by ChatGPT were more open-ended than traditional skill practice, and that the activities allowed students to explore concepts.
4. Engaging/relatable	In this theme, the teachers accentuated that the tasks generated by ChatGPT would be engaging and fun for students, and that they related to the real-world and students’ lives.
5. Adjustable/Differentiated	In this theme, the teachers recognized that the ChatGPT activities were designed to differentiate learning to individual students’ learning needs and backgrounds.
6. Multiple representations	In this theme, the teachers appreciated that the activities created by ChatGPT incorporated multiple representations of mathematical concepts – like tables, equations, graphs, and real-world contexts.

Teachers also had critiques of the GenAI output that they gave when asked what they noticed and wondered (Table 2). For example, one teacher thought the active learning activities were *not* engaging: “This activity lacks fun, and students may feel bored.” There were regular comments from teachers about the lack of fit of what the AI outputted to their students and context; one teacher said, “That’s not what that [mathematical equation] looks like for an 8th grader. That’s ridiculous.” Another described how “This one took me forever to get to an example that I liked. I

had to get really specific with the student’s interests because originally, it just tried to make the original problem into a full-on theater production.”

One teacher commented specifically on the mathematical accuracy of ChatGPT’s output: “So I already see mathematical flaws with this. And this is why... I wouldn't ever rely on something like this because... I would have to fix it. So it's easier for me to just create it.” Both this quote and the quote before it capture another theme – that teachers could find the GenAI to be too time-consuming or not efficient for their work process. Teachers also emphasized that they wanted ChatGPT to directly create student-facing multimedia materials like worksheets and PowerPoints rather than just provide lesson plans, with one teacher saying, “Yeah, that would be great [if ChatGPT could create] not just worksheets, but maybe the index cards it was referring to, or whatever. That would be really nice.” Finally, two teachers were concerned when the ChatGPT lesson plan required materials they did not have or that would be overly time-consuming to construct themselves, especially if the materials needed to last through repeated use by their students.

**Table 2. Teachers’ reflections on negative elements of using ChatGPT to enhance generative learning**

<b>Theme</b>	<b>Description</b>
1. Fit to students	For this theme, teachers critiqued the fit of the ChatGPT materials to their students, including developmental appropriateness.
2. Boring	For this theme, teachers would describe the ChatGPT activities as lacking in their ability to engage students in learning.
3. Preparation time	For this theme, teachers would describe how actually being ready to implement the activities that ChatGPT gave would be time consuming for them in terms of preparation time, and/or that it may be faster for them to just design the activities themselves.
4. Lack of handouts	For this theme, the teachers would express concerns about the lack of student facing materials like handouts or flash cards to accompany the lesson plan being generated by ChatGPT
5. Mathematical flaws	For this theme, teachers would question the mathematical accuracy and rigor reflected in the tasks and lessons generated by ChatGPT.
6. Lesson materials	For this theme, teachers would describe how the ChatGPT lesson required materials that they did not have or that many other teachers would not have, and/or materials that would be time-consuming or impossible for them to make by hand.

When teachers were subsequently asked what developers should keep in mind when creating AI tools for active learning, many repeated a negative element they had cited in Table 2 instead of framing it in a more constructive way. For example, one teacher described how AI systems should give “ideas for learning activities that we can implement in 5-10 mins” (corresponding to Theme 3 in Table 2), while another suggested a system with “With sound [mathematically accurate] material and then student-facing worksheets... that you can manipulate” (corresponding to Themes 4 and 5 in Table 2). Teachers also brought up issues such as the length of ChatGPT’s typical output (corresponding to Theme 3 in Table 2): “Because, just scrolling

through here, by the time you get to the bottom, you've forgotten what was up here at the top. It could be, I don't know, maybe even tabbed?" A second teacher also described the appearance of the output, along with its content: "I always want my materials to be presentation-ready. So, are they ready for publication? Are they using math type? Is it using correct academic language?" This quote corresponds to Themes 3, 4, and 5 in Table 2.

Another teacher called attention to how flexible the materials need to be for different use cases, referencing Theme 1 in Table 2: "You have students that don't speak English. You have some teachers that don't know the material. You have some students who have gaps." However, a different teacher challenged the necessity of this level of customization of the learning materials, arguing that they would not necessarily want to give the LLM fine-grained information about their students, referencing Theme 3 in Table 2: "The model should know something about what works already. I don't have time to customize. I understand that we should. I understand that I should teach every student differently. I understand that would be awesome. But I got 150 students. I can't do that."

The teachers who were interviewed were also asked about challenges relating to using GenAI in this manner. They brought up big-picture issues not reflected in Tables 1-2, such as issues of surveillance, reliance on AI and AI as cheating, teacher training on using AI, and the de-professionalization and disconnection of teachers through AI. One teacher described how "I noticed you don't have to think for yourself anymore. This is scary that it [ChatGPT] did that; that's scary that [ChatGPT] can give you all of this." Another described how "There has to be teacher training... And you know I hate to say, teacher-proof, right?" Another described how AI was trained using videos of teachers, such that it could replace teachers:

*And this way they train all the AI... I think they are pretty successful doing all this research, but I hate it because of that. Because when they succeed, the teacher has no job, and they could use all the AI to teach online high schools and online universities. They don't need all the faculty anymore. But you know, that's the goal for these engineers. They think teaching is a science. I disagree with them. Teaching is a part of [the] arts and is very important of human feelings and emotional connections.*

Finally, one teacher described how "I think it's wonderful to have resources. I mean, the reason we go to conferences, and we network, and we're involved in teacher communities, is to share materials... I'm always getting emails saying, 'Do you have anything for this? Do you have anything for that?'" This implies that with GenAI tools like a tool that incorporates active or generative learning, teachers may have less reason to network and learn from each other.

### ***RQ2: GenAI as a Brainstorming Partner for Teachers Posing Questions***

Our second research question examines teachers' reflections when they were using ChatGPT as a support to complete a generative learning task – posing mathematical questions aligned to their grade level standards from real-world photos of their communities and surroundings (Figure 2). Teachers recognized a variety of strengths and weaknesses of GenAI as a brainstorming partner for generative learning, as shown in Tables 3-4. In terms of affordances (Table 3), teachers recognized that brainstorming with GenAI could enhance their creativity and give them many

varied examples to choose from, with one teacher describing how “When you are not inspired, we could use this tool to help us be creative.” The teachers also appreciated that you could go back and forth with ChatGPT through iterative prompting and that it was a flexible and versatile tool: “You can build on previous requests and change out and outcomes to the slightest or highest degree.” Teachers also commented that ChatGPT generates responses quickly and that it has the potential to save them time in lesson planning. One teacher also identified how ChatGPT was able to make his mathematical questions more specific: “My problem statements focused vaguely on concepts, but ChatGPT brought in actual numbers that transformed my problems from conceptual ones to calculable ones.”

**Table 3. Teachers’ positive reflections on using ChatGPT to generate questions from photos**

<b>Theme</b>	<b>Description</b>
1. Creativity	For this theme, teachers describe how ChatGPT enhanced their creativity at question-asking, incorporating ideas into the brainstorming process that were different from their own.
2. Multiple examples	For this theme, teachers discuss how ChatGPT can generate both a large quantity and large variety of questions from photos.
3. Iterative prompting	For this theme, teachers recognize the affordance of being able to engage in prompting and iteration with ChatGPT, allowing it to refine its output to fit their needs.
4. Time efficient	For this theme, teachers describe how using ChatGPT allows ideas to be generated in a time efficient manner, in real time as they have questions, saving time in lesson preparation.
5. Generates details	For this theme, teachers appreciate that ChatGPT could take a picture and layer onto that picture both specific numbers and details not in the picture about a hypothetical mathematical situation.
6. Versatile/flexible	For this theme, teachers appreciate that ChatGPT offered ideas for questions that were both versatile and flexible, allowing them to meet their students’ needs.

In terms of concerns (Table 4), teachers again discussed the fit of the generated content to their students, mathematical needs, and context: “It doesn’t have a good grasp of what mathematical concepts are appropriate for a specific grade level. I asked it for Algebra 2-level math questions, and it gave me a lot of geometry-related questions.” Another teacher wrote, “Sometimes the answer is theoretical; we have to decide whether it is appropriate to use in the classroom.” Teachers also commented on issues with hallucinations and the lack of reliability of ChatGPT, with one teacher describing how “GPT could be wrong. Human[s] need to justify whether their answer is right or wrong.” One teacher again emphasized the costs of using ChatGPT in terms of teachers’ time, describing how she was “Spending too much time on figuring out what to ask instead of actually focusing on the task on hand.” Finally, one teacher also commented on its reliability and felt that the questions generated by ChatGPT were not answerable, “It doesn’t give enough context for me to solve the problems, which means my students will also struggle to solve the problems. “

**Table 4. Teachers' negative reflections on using ChatGPT to generate questions from photos**

<b>Theme</b>	<b>Description</b>
Grade level misalignment	For this theme, the teachers would critique the questions ChatGPT generated as not aligned to their grade level.
Math topic misalignment	For this theme, the teachers would critique the ChatGPT-generated questions as not aligned to the mathematical topic they specified.
Sensitivity to prompts	For this theme, the teachers expressed concern that ChatGPT struggled to understand what they actually meant when they prompted it.
Not reliable	For this theme, teachers described how they could not rely on ChatGPT's output, citing issues like hallucinations.
Not time efficient	For this theme, the teachers described how using ChatGPT as a brainstorming partner was too time-consuming for them.

## **Discussion**

Our study examined the potential of generative AI to support mathematics teachers in creating lesson plans that incorporate generative learning approaches, as well as to assist with the generative learning task of generating mathematical questions from real-world photos. We first discuss results for lesson planning in general, then for generative learning specifically, and then explore bigger-picture issues teachers face with GenAI.

In its current form, ChatGPT seemed to be almost, but not quite, useful as a lesson planning tool. On the surface, its outputs were detailed, and its lessons were engaging. From a quick read, ChatGPT's capabilities would come across as very impressive. However, teachers noticed a variety of issues upon closer examination. ChatGPT's output did not reflect a pedagogical understanding of different math standards as they are implemented at different grade levels, which made the lesson ideas seem like a poor fit to teachers. Teachers consistently pointed out that student-facing materials like worksheets and slides were not easily generated and that issues with reliability and AI's tendency to hallucinate gave teachers pause on whether using ChatGPT was worth it for them. This is consistent with other research that has shown limitations of ChatGPT for lesson planning in mathematics education (e.g., Aga, Sawyer, and Wolfe 2024; Kim, Park, and Joung 2024; Kuusemets et al. 2024). Overall, the lesson plans seemed too general or generic to be used in practice in an immediately actionable way, which is what the teachers wanted and needed. Some of these issues may be ameliorated by teachers gaining expertise with prompting and iterative revision of ChatGPT materials, as is suggested by prior research (Biton and Segal 2025; Broutin 2024). Other issues may be endemic to the AI itself and may not be easily solved through new prompting strategies.

When considering the possibilities of GenAI as a lesson planning and brainstorming tool for generative learning activities (Brod 2021) specifically, new considerations arise. First, ChatGPT's output seems to generally give reasonable ideas that are consistent with principles of generative learning (Fiorella and Mayer 2016), although they are somewhat generic and vague.

ChatGPT integrates a lot of attractive qualities into its generative learning lesson planning that teachers appreciate – such as hands-on learning, real-world examples, exploration, and multiple representations. Also, importantly, teachers seem to realize that ChatGPT generates things they would not be able to think of on their own, which expands their horizon of possibilities as teachers and designers, which is especially important for teachers looking for creative ways to start integrating in more generative learning into their courses. However, generative learning is a complex construct that can involve a wide variety of different approaches (e.g., Brod 2021; Fiorella and Mayer 2016), and perhaps because of this, the outputs given by ChatGPT for implementing generative learning tended to be long and incorporate a lot of different generative learning strategies. It can be time-consuming for teachers to figure out how to integrate these ideas for generative learning into their instruction and to adapt them to suit their context and needs, especially since ChatGPT seems to give somewhat generic responses (see also Aga, Sawyer, and Wolfe 2024), at least initially.

Finally, we found that teachers identify serious big-picture issues with the use of generative AI for planning lessons involving generative learning that we cannot ignore. They are concerned about surveillance of themselves as it relates to increasingly adaptive suggestions for and iterations of lessons they request. They also worry about teachers becoming over-reliant on AI and not having to think for themselves, citing that generative AI might serve to reduce teacher-teacher collaboration and networking. Finally, they are concerned with issues of teachers getting appropriate training to use GenAI effectively and appropriately, and they are concerned with the idea that GenAI could take away their jobs as teachers. Indeed, research suggests teachers spend around 50% of their work time doing daily non-instructional tasks like lesson planning, grading, and communicating with parents (Cardona, Rodríguez, and Ishmael, 2023). If GenAI is able to reduce this time by half, will this mean that teachers get more breathing room and time for professional activities, or will it simply mean that we will hire 25% fewer teachers? GenAI is not yet at this point, but this will be an important consideration in the future.

### **Policy Action Recommendations for Local Adoption and Implementation**

We close with five actionable recommendations for teachers using generative AI for lesson planning for generative learning. First, our results suggest that we need AI systems for teachers that are designed to **learn and adapt to teachers' specific and individual preferences** for implementing generative learning – some teachers value precise mathematical language aligned with their instructional approach, others value student engagement, exploration, and hands-on learning, while still others have limitations on the materials they will be able to provide for lessons. All teachers strongly want output that is well-aligned to their particular grade level, the content area and curriculum they are teaching, how they individually prefer to teach it, and the mathematics standards for which they are held accountable. This focus on specificity to teachers' preferences needs to, however, be balanced with issues of surveillance of teachers using GenAI.

Second, we need integrated GenAI programs for generative learning that will **generate a variety of multimedia student materials to go with lesson plans** – including worksheets, PowerPoints, handouts and templates, and mathematical visuals. These programs should have built-in ways to have teachers still collaborate and share their experiences using these materials, as well as allowing teachers to build on, adapt, and reflect on materials used by other teachers. Third, we

need to find ways to have GenAI give suggestions for generative learning that involve **shorter, more “bite-sized” chunks of active learning activities for teachers to implement** – perhaps just involving one or two active learning approaches at a time instead of many, in activities that can be more easily integrated into a full instructional day. Teachers will have more success making small modifications to their current practice to incorporate more generative learning rather than completely overhauling every element of their instruction. This activity generation with GenAI should be coupled with sustained teacher professional development; this should include (a) helping AI novices build an accurate mental model of how LLMs work (e.g., thinking of AI as closer to predictive text than ‘intelligence’), (b) cautioning about the dangers of AI use, including bias (Bender et al. 2021) and hallucinations, and how to identify and counteract these risks (c) training on prompt engineering to ensure high-quality outputs, (d) implementation training on how to incorporate the outputs into classroom instruction effectively, (e) an ongoing community of practice for teachers to periodically share successes and failures in using AI to facilitate active learning in their classroom.

Fourth, **we need research that explicitly addresses issues of time efficiency with using GenAI for lesson planning.** We need to be able to answer questions like, how much time does GenAI save teachers? Are the student-level customizations that GenAI proposes realistic for implementation? How can GenAI be used in a time-efficient way? What impact does this time efficiency have on students’ interest and learning? In our own work (Walkington et al. under review), we are having teachers adapt mathematics problems to be personalized to students’ interests using GenAI. We are measuring how much time it takes teachers to formulate each problem with GenAI and will compare this to formulating problems without GenAI. We are also looking at whether spending more time formulating problems with GenAI predicts students being more interested in a problem and learning more from solving a problem (see Walkington et al. under review) and which specific teacher prompting moves with GenAI predict positive outcomes for students. This kind of research will be key in determining when GenAI is ready to be scaled as a tool for instructional planning. We also need research on where this saved time, if there is saved time, should go, so that it does not end up simply costing teachers their jobs.

Fifth, as a field, **we need to address teachers’ big-picture concerns about generative AI,** including issues of surveillance and privacy, AI replacing teachers, over-reliance leading to lack of teacher collaboration, and appropriate teacher training. These concerns can be addressed through thoughtful implementation of policies and funding for teachers to engage with AI in particular ways. We need to improve teacher education and training on AI so that teachers understand how GenAI works, its affordances and limitations, and its ethical implications, such that teachers feel comfortable living in a world where GenAI is used in education and feel confident in being able to handle the myriad of issues that arise with its use.

**Acknowledgements:** This research received funding from Rice University – OpenStax, via the Bill & Melinda Gates Foundation.

## References

- Aga, Z. G., A. G. Sawyer, and M. Wolfe. 2024. "More Confidence More Critical: Investigating Secondary Mathematics Preservice Teachers' Use of an Artificial Intelligence Chatbot as a Curriculum Development Tool." *American Journal of Creative Education* 7 (2): 63–78.
- Bender, E. M., T. Gebru, A. McMillan-Major, and S. Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜." In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, edited by M. Elish, W. Isaac, and R. Zemel, 610–23. <https://doi.org/10.1145/3442188.3445922>.
- Biton, Y., and R. Segal. 2025. "Learning to Craft and Critically Evaluate Prompts: The Role of Generative AI (ChatGPT) in Enhancing Pre-Service Mathematics Teachers' TPACK and Problem-Posing Skills." *International Journal of Education in Mathematics, Science, and Technology (IJEMST)* 13 (1): 202–23.
- Braun, V., and V. Clarke. 2012. *Thematic Analysis*. American Psychological Association.
- Braun, V., and V. Clarke. 2019. "Reflecting on Reflexive Thematic Analysis." *Qualitative Research in Sport, Exercise and Health* 11 (4): 589–97.
- Brod, G. 2021. "Generative Learning: Which Strategies for What Age?" *Educational Psychology Review* 33 (4): 1295–1318. <https://doi.org/10.1007/s10648-020-09571-9>.
- Broutin, M. S. T. 2024. "Exploring Mathematics Teacher Candidates' Instrumentation Process of Generative Artificial Intelligence for Developing Lesson Plans." *Yükseköğretim Dergisi* 14 (1): 165–76. <https://doi.org/10.53478/yuksekogretim.1347061>.
- Cardona, M. A., R. J. Rodríguez, and K. Ishmael. 2023. *Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations*. Office of Educational Technology, U.S. Department of Education. <https://policycommons.net/artifacts/3854312/ai-report/4660267/>.
- Fiorella, L., and R. E. Mayer. 2016. "Eight Ways to Promote Generative Learning." *Educational Psychology Review* 28: 717–41. <https://doi.org/10.1007/s10648-015-9348-9>.
- Fiorella, L., and R. E. Mayer. 2020. *Learning as a Generative Activity: Eight Learning Strategies That Promote Understanding*. Cambridge University Press.
- Gómez Marchant, C., and H. Hardison. 2024. "In the Shadows of Burgeoning Colossi: The Whiteness of AI in Mathematics Teacher Education." *AMTE Connections*. <https://amte.net/connections/2024/05/connections-thematic-articles-artificial-intelligence-mathematics-teacher>.
- Google Gemini. n.d. *Gemini*. <http://gemini.google.com>.

- Khan Academy. n.d. *Khanmigo by Khan Academy*. <https://www.khanmigo.ai/>.
- Kim, Y., M. Park, and E. Joung. 2025. "Exploring the Integration of Artificial Intelligence in Math Education: Preservice Teachers' Experiences and Reflections on Problem-Posing Activities with ChatGPT." *School Science and Mathematics*, 1–15. <https://doi.org/10.1111/ssm.18336>.
- Kuusemets, L., K. Parve, K. Ain, and T. Kraav. 2024. "Assessing AI-Generated (GPT-4) Versus Human-Created MCQs in Mathematics Education: A Comparative Inquiry into Vector Topics." *International Journal of Education in Mathematics, Science, and Technology (IJEMST)* 12 (6): 1538–58. <https://doi.org/10.46328/ijemst.4440>.
- MagicSchool.ai. n.d. *Magic School: AI Built for Schools*. <http://magicschool.ai>.
- Mayer, R. E. 2021. *Multimedia Learning*. 3rd ed. Cambridge University Press.
- Michael, K., M. Alemu, Y. Desie, M. Atnafu, S. Assefa, C. Regassa, H. Wodaj, and A. Abate. 2023. "Understanding and Practice of Active Learning Among Upper Primary School Science and Mathematics Teachers." *Heliyon* 9 (6): e16854. <https://doi.org/10.1016/j.heliyon.2023.e16854>.
- RAND. 2024. "Survey of American Math Educators." *Unpublished Raw Data*. Retrieved from <https://danmeyer.substack.com/> (accessed October 8, 2024).
- Sawyer, A., and Z. Aga. 2024. "Counterexamples to Demonstrate Artificial Intelligence Chatbot's Lack of Knowledge in the Mathematics Education Classroom." *AMTE Connections*. <https://amte.net/connections/2024/05/connections-thematic-articles-artificial-intelligence-mathematics-teacher>.
- Wachira, P. 2025. "Mathematics Teaching and Assessment in the Age of Generative AI." In *Educational Assessments in the Age of Generative AI*, edited by P. Wachira, X. Liu, and S. Koc, 315–42. IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-6351-5.ch011>.
- Yen, A. Z., and W. L. Hsu. 2023. "Three Questions Concerning the Use of Large Language Models to Facilitate Mathematics Learning." *arXiv Preprint arXiv:2310.13615*. <https://doi.org/10.48550/arXiv.2310.13615>.