Mathematics Teachers Using Generative AI to Personalize Instruction to Students' Interests

Theodora Beauchamp and Candace Walkington (Southern Methodist University)

Introduction

Effectively leveraging Generative AI (GenAI) in the future will be a combined effort to develop greater teacher content knowledge alongside teachers' continued growth in understanding the latest AI tools, creating challenges for teacher education (National Council of Teachers Mathematics, 2024). Teachers find that GenAI can automate administrative tasks, personalize learning experiences for students, and provide analysis of student data. GenAI can increase teachers' efficiency, allowing students access to more content because teachers do not have to create a curriculum. Magic School (magicschool.ai) is a GenAI program designed specifically for educational purposes, with teacher- and student-facing products. Here, we focus on the educator tools in Magic School that allow teachers to personalize learning according to their students' interests (Walkington, 2013) which is an asset-based approach to improving learning using AI (Ocumpaugh et al., 2024). Here we will describe lessons learned as teachers engage in personalization using Magic School, which is currently the most-used educator AI platform, used by more than a million teachers.

Background

GenAI can increase the relevance of school-based tasks for learners. When learners find tasks they are completing relevant, this can trigger their interest (Hidi & Renninger, 2006) in their learning. This triggered interest can lead to a variety of positive effects on the learning process and outcomes (Renninger & Hidi, 2015). Research on interventions where students learning mathematics solve or create their personalized problems with GenAI has found mixed results. A study of college students found that career-personalized mathematics problems written by GenAI were sometimes appreciated by students but that there were concerns about adding additional lengthy contextual details to problems (Einarsson et al., 2024). A study of middle school students writing personalized math problems with GenAI suggests the efficiency and relevance of this approach, but problems lacked deeper levels of authenticity (Walkington et al., 2024).

Kaplan-Rakowski et al. (2023) distributed a validated survey to teachers that measured their perceptions of GenAI's benefits and shortcomings when used in the classroom. They found that teachers were likely to want to integrate GenAI into their teaching, expressing positive sentiments. Considering the ethical implications of using GenAI in the classroom is also key. There are well-documented biases in large language models that harm marginalized groups (Bender et al., 2021). Aguilar (2024) conducted a study of how K-12 teachers think about ethics with respect to GenAI in the classroom. He found that "while some were staunch proponents of strict ethical guidelines, others believed in a more outcome-driven approach" (p. 25).

Context of Data

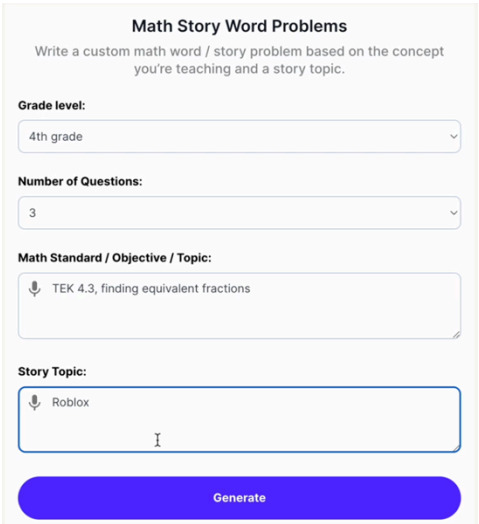
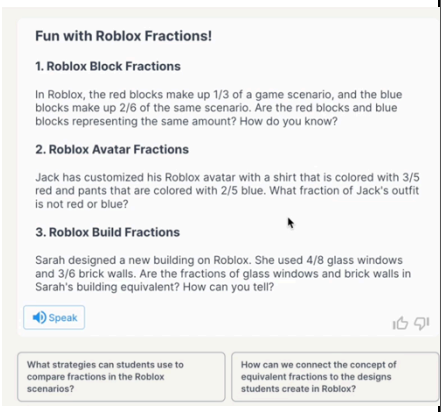
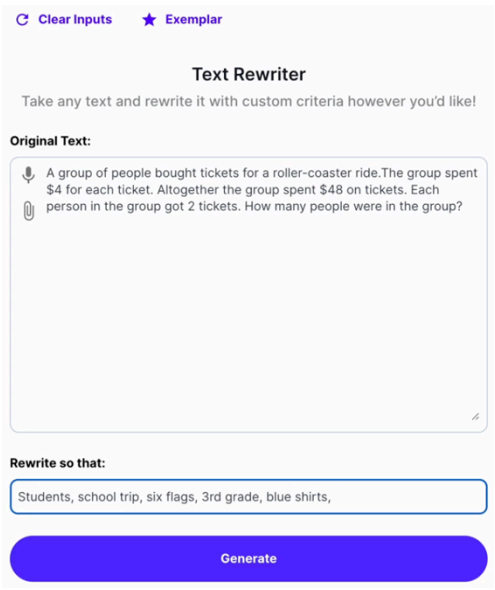
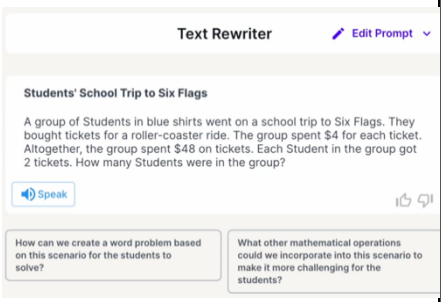
We discuss 12 teachers (9 female, 3 male) in a university teacher education class focusing on elementary math methods; two participants identified as Asian, one as Chinese, two as Black, two as Hispanic or Latino, one as Spanish, and five as White. Three were pre-service teachers, eight were first-year teachers, and one was a second-year teacher. The in-service teachers spanned various grade levels: one teaches Kindergarten, one Second grade, one Fourth grade, two Fifth grade, two Sixth grade, one Seventh grade, and one teaches Algebra 1.

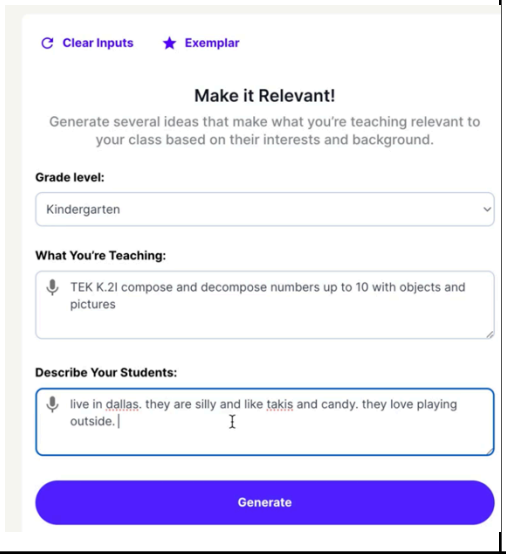
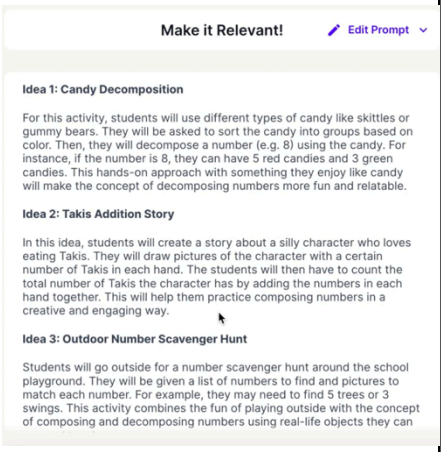
The teachers used three of Magic School's GenAI tools to personalize learning: *Math Word Problem Maker*, *Make it Relevant,* and *Rewrite* (Figure 1). They were given the choice to use math word problems from their state standardized test or others they found from district curriculum. Participants shared the grade level they teach and information about their classroom to see what Magic School generated regarding ideas for activities. Teacher's screen-recorded their usage of each tool while collaborating with a partner. They used Google Slides to share their products with others. The session closed with a recorded whole-group discussion and teachers completing a short survey about the AI tools. This study examines participants' entries into the AI tool, what was produced by the teachers through the AI tools, participants' interactions with each other through screen recordings, participants survey responses, and output generated by the AI tool.

Examples of Math Teachers Using or Discussing Magic School

Figure 1 shows three different AI tools in Magic School (left column), an example of the output generated by the AI tool based on participants' prompts (middle column), and participant responses to post-survey questions asking them about their experiences (right column). The survey responses, in the context of the way in which participants interacted with the tools, were used to extract themes relating to participants' engagement with AI tools. Four key themes emerged: issues of bias, efficiency, knowledge of student interests, and student motivation (Figure 1).

Figure 1: Noticing's by participants while using Magic School AI tools

Math Story Word Problems Tool	GenAI Output	Survey Response *[Theme]*
		"Some issues with not understanding Roblox and how it works, gave good examples and variety of problems when prompted. The framing of the questions are a little similar." (First-year 6th grade teacher) [*Knowledge of Student Interests*] *[Efficiency]*
Text Rewriter Tool	**GenAI Output**	**Survey Response**
		"Positives: regenerated problems to make problems more relevant and engaging for students. Negatives: we typed in a lot of edits, so the AI didn't incorporate all the edits we suggested." (First-year 5th grade teacher) *[Student Motivation]*

Make it Relevant Tool	GenAI Output	Survey Response
		"Very creative! Prompts were too long at times." (First-year Kindergarten teacher) *[Student Motivation]*

The teachers noted that issues related to *bias* were present; a first-year 7th-grade female teacher described how: "We asked it to rewrite a math problem, about RAEL [Recently Arrived English Learner] student, so like recently arrived immigrant who did not speak like English, who has never spoken English much before from Bolivia. We were like, hey, rewrite this problem to be about someone from Bolivia, and literally, the only thing it changed was that it said like in the title of the word problem. The students went to a Bolivian amusement [park]." Other teachers also noted GenAI's *efficiency* for lesson planning. A female master's student who was student-teaching Algebra 1 stated, "It was very efficient for most of the problems, but I feel like we just got there. There might be a limit depending on how the problem is worded." Teachers also commented on student motivation. A first year, male, 6th-grade teacher shared, "It came up with a wide variety of ideas, especially when it came to higher education and exploring careers. TikTok activity is also pretty good and thoughtful. Gives a good baseline," Finally, they discussed the AI's knowledge of student interests. A first-year female 5th-grade teacher said, "It (Make it Relevant) was really good for making it specific towards our students and their interests."

Lessons Learned

GenAI, in combination with prompt engineering, can be a powerful tool for teachers. We found that when teachers used these tools, teachers' knowledge of students, student motivation, potential bias, and efficiency were important noticing's by the teachers.

As GenAI becomes more robust in its offerings, the need for teachers' understanding of its usage remains vital to prepare teachers for ways in which it might be used in classrooms and students' future careers. As described above, our

participants identified both positive and negative elements of their experience, with one first-year male 5th grade teacher saying: Positives: it took the prompts that I took to make relevant and made challenging problems. Negatives: instead of searching for and typing in TEKS (Texas Essential Knowledge and Skills; the state standards for Texas), we wish we had the ability to see the TEKs on the actual site."

This study did bring some new insights. Magic School's Word Problems tool allowed for different personalized problems to be created, which used a specific mathematics state standard and a personalized story problem topic from the teachers. Teachers noted that having a variety of word problems generated was a positive outcome: "I love what AI was able to do with such specific topics like makeup and TikTok!", commented a pre-service, female, undergraduate student.

The teachers noted limitations when using the Rewrite It tool in Magic School. Several teachers noted that it did not reflect their student's background, ethnicity, and interests in a personalized way. One female, first-year, Algebra 1 teacher remarked, "It just did not do that much making it relevant, like just common things like changing the words but not much of the problem structure." It mentioned attributes of the students, but the tool did not give responses indicating it understood the topic in depth. Lastly, the Make it Relevant tool was also noted to have positive components as it gave teachers multiple ideas for activities generated by state mathematics standards and more importantly, by a teacher's description of their student's interests.

Studying Magic School within a teacher education program enabled an understanding of teacher perspectives and how GenAI tools for math education could be used for personalized instruction. This study had some important limitations, including the small sample size of teachers, the fact that the teachers were inexperienced in teaching, and that we only examined three different AI tools. More research should be done to look at how teachers engage with generative AI tools intended to support their instruction and the impact that generative AI enhanced story problems might have on mathematics achievement. GenAI will become increasingly used by teachers as its capabilities increase and it becomes integrated into teacher training and professional development, and our understanding of how to best support teachers in using these tools must also continue to develop.

Acknowledgements

This work was made possible through the support of the Learning Engineering Virtual Institute (LEVI), a sponsored project of Rockefeller Philanthropy Advisors.

References

Aguilar, S. (2024). Ethics in Generative AI: Report from the Field. *In Critical Thinking and Ethics in the Age of Generative AI in Education* (pp. 21-26). USC Center for Generative AI and Society. https://rossier.usc.edu/documents/usc-center-generative-ai-and-society-report

Einarsson, H., Lund, S. H., & Jónsdóttir, A. H. (in press). Application of ChatGPT for automated problem reframing across academic domains. *Computers and Education: Artificial Intelligence*, 100194.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational psychologist*, *41*(2), 111-127.

Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research*, *34*(2), 313-338.

National Council of Teachers of Mathematics. (2024). *Artificial Intelligence and Mathematics Teaching, A Position of the National Council of Teachers of Mathematics [Position Statement].* https://www.nctm.org/standards-and-positions/Position-Statements/Artificial-Intelligence-and-Mathematics-Teaching/

Ocumpaugh, J., Roscoe, R. D., Baker, R. S., Hutt, S., & Aguilar, S. J. (2024). Toward Asset-based Instruction and Assessment in Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, 1-40.

Renninger, K. A., & Hidi, S. (2015). *The power of interest for motivation and engagement*. Routledge.

Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of educational psychology*, *105*(4), 932.

Walkington, C., Milton, S., Pando, M., Lipsmeyer, L., Sager, M., & Beauchamp, T. (2024). Adolescents Using Generative AI to Engage in Mathematical Problem-Posing. Paper to be presented at the *International Congress on Mathematical Education (ICME-15)*. Sydney, Australia.