

Using ChatGPT As A Thought Partner in Writing Relevant Proportional Reasoning Word Problems

Andrea Berryhill (Mississippi State University), Lendon Chandler (Mississippi State University), Liza Bondurant (Mississippi State University), and Bima Sapkota (University of Texas, Rio Grande University)

Introduction

The integration of technology into education has ushered in a wave of innovative teaching and learning approaches. Among these advancements, large language models (LLMs) like ChatGPT have gained considerable attention for their capacity to produce high-quality text, translate languages, and respond to open-ended queries. Educators, particularly mathematics teacher educators (MTEs), are intrigued by the potential of LLMs to enhance various facets of teaching, including lesson planning, although they recognize the need for cautious exploration and critical reflection in this emerging landscape (NCTM, 2024; Meyer, 2023; Sapkota & Bondurant, 2024).

Despite the abundance of online and print mathematics education resources, many students lack access to high-quality instructional materials (HQIMs). Mathematics teachers (MTs) often grapple with the overwhelming array of available resources and may struggle to discern which ones align with HQIM criteria. Consequently, well-intentioned MTs may resort to using personal funds to procure materials that fall short of meeting HQIM standards (TNTP, 2018).

ChatGPT has emerged as a promising tool to assist MTs in lesson planning (Staffel, 2023; Trust et al., 2023; Yu, 2023). MTs can use ChatGPT to adjust application-based problems to student interests (National Council of Teachers of Mathematics [NCTM], 2024). However, researchers have identified limitations, including the generation of inaccurate or nonsensical outputs, an inability to comprehend or build upon incorrect inputs, and a failure to meet the humanistic "caring test" (NCTM, 2024; Meyer, 2023; Sapkota & Bondurant, 2024). In this article, two preservice teachers (PSTs) and two MTEs, document our collaborative exploration into leveraging ChatGPT as a thought partner in crafting relevant proportional reasoning word problems. We detail our iterative process of refining inputs and critically evaluating ChatGPT outputs. The insights gleaned from our study can inform educators' strategic utilization of LLMs as thought partners in lesson planning.

Relevant Proportional Reasoning Word Problems

In contrast to the beliefs of some practitioners and researchers in mathematics education, students have demonstrated greater success in solving word problems that are relevant to their real lives compared to mathematically equivalent equations (Koedinger & Nathan, 2004). Discussion about abstract versus contextualized, "pure" or "applied" mathematics sometimes overlooks the core issue, student engagement. Providing real-life context that is relevant and affirming to each and every student can make mathematics more accessible to a broader range of students (Aguirre, 2015;

Grandin, 2022; Ladson-Billings, 2022; Marshall, 2023; Meyer, 2010, 2015). This research motivated our desire to create word problems that are relevant to the students we work with. Regarding the content, we decided to focus on relevant word problems that require students to identify the constant of proportionality because many U.S. struggle with this skill, which is a major concept in number (National Assessment of Educational Progress [NAEP], 2022; NCTM, 2020).

Iterative Refinement of Inputs and Critical Examination of Outputs

We began by discussing the key terms we should include in our ChatGPT inputs. We decided to use the key terms “relevant,” “constant of proportionality,” “proportional relationship,” “Mississippi (MS),” “middle school,” and “word problem.”

The example 1 input and output, found in Table 1, is representative of our initial inputs and outputs. Next, we will share the strengths, weaknesses, and revisions we made.

Table 1:

Example 1 ChatGPT Input and Output

Input	Output
Write a relevant word problem for middle school students in Mississippi that requires them to identify the constant of proportionality.	In a local Mississippi farmer's market, Julia is selling baskets of fresh peaches. The cost of the peaches is directly proportional to the weight of the peaches in each basket. Julia charges \$2.50 for a basket that weighs 2 pounds. If Sarah wants to buy a basket that weighs 4 pounds, how much would she need to pay?

The strengths of example 1 included:

- agriculture was once the #1 industry in Mississippi (MS),
- farmers' markets are common in MS,
- the problem involves a female character which female students may relate to, and
- the problem involves a proportion.

However, the weaknesses of example 1 included:

- agriculture is currently only 17% of MS GDP (IBIS World, 2023),
- peaches are not a common crop in MS (MDAC, 2023),
- the current price of one pound of peaches is \$2.99 (Kroger, 2023),
- the question does not ask for the constant of proportionality, also known as the unit rate,
- the wording sounds like Julia paying for weight of basket not of peaches, and
- Julia is not a common name for girls in MS (MSDH, 2023).

We revised the ChatGPT output to include the #1 crop in MS, eggs (MDAC, 2023), the current price a dozen organic large eggs, \$4.49 (Kroger, 2023), the #1 name for a girl in MS, Ava (MSDH, 2022), removed mention of the weight of the container holding the crop, and focused the question on the constant of proportionality.

Based on ChatGPT’s initial outputs, we made two major changes to our inputs. First, to ensure ChatGPT generated relevant outputs, we added specificity to our inputs. Instead of simply requesting “relevant” word problems, the team provided specific contexts they wanted to incorporate. In example 2, we specified the context of an MS recipe. The second step involved targeting specific learning objectives. We removed the key term “proportional relationship” to ensure the outputs targeted the skill of finding the constant of proportionality also known as the unit rate.

The example 2 input and output, found in Table 2, is representative of the inputs and outputs after our revisions.

Table 2:
Example 2 ChatGPT Input and Output

Input	Output
Write a relevant word problem for middle school students in Mississippi that requires them to find the constant of proportionality of potatoes in a Mississippi recipe for potato soup and compare it to the constant of proportionality of potatoes in an Irish recipe for potato soup.	In Mississippi, a family recipe for potato soup calls for 4 pounds of potatoes to serve 6 people. Meanwhile, in Ireland, an Irish recipe for potato soup requires 6 pounds of potatoes to serve 8 people. Find the constant of proportionality for the amount of potatoes needed in each recipe. How might the context influence the amount of potatoes needed per person in each potato soup recipe?

The strengths of example 2 included:

- the MS context provides a mirror (Bishop, 1990; Gutiérrez, 2018) as potato soup is popular in MS (Kimbrough, 2021), and the Ireland context provides a window and potatoes are a staple in the Irish diet (Pope, 2020),
- the problem required students to find two different constants of proportionality and then compare them,
- the quantities in the recipes were accurate, and
- it is reasonable that the MS recipe has a lower constant of proportionality ($\frac{2}{3}$ pounds per person) of potatoes than the Irish recipe ($\frac{3}{4}$ pounds per person) because potatoes are considered a staple in Irish diets (Pope, 2020).

We did not see any major flaws in the example 2 output, so no revisions were needed.

Lessons Learned and Implications for Practice

We gained valuable insights that can inform educators seeking to leverage the potential of LLMs in their lesson-planning endeavors. We had three key takeaways. For one, LLMs are wonderful thought partners, but a human-in-the-loop is needed. LLMs are valuable tools for generating ideas, but educators must remain the architects of the learning experience, exercising critical judgment and pedagogical expertise to refine and adapt the outputs (Sapkota & Bondurant, in press) Secondly, we realized the need

for iterative revisions of our inputs. Effective LLM integration hinges on the ability to craft clear, specific, and strategic inputs that guide the LLM toward the desired outputs. Finally, we discovered the value of collaboratively critiquing and revising the inputs and outputs. The different perspectives of each of our team members helped us critically examine the outputs and creatively engineer the inputs and revisions. Our takeaways align with the Position of the National Council of Teachers of Mathematics on Artificial Intelligence and Mathematics Teaching (2024).

References

- Aguirre, J. M. (2015). *Enhancing the common core with culturally responsive mathematics teaching: Key principles and strategies*. National Council of Teachers of Mathematics
- Bishop R. (1990). Mirrors, windows, and sliding glass doors. In *Perspectives: Choosing and using books for the classroom* (Vo. 6, No. 3). Ohio State University. <https://scenicregional.org/wp-content/uploads/2017/08/Mirrors-Windows-and-Sliding-Glass-Doors.pdf>
- Grandin, T. (2022). Against Algebra. *The Atlantic*. <https://www.theatlantic.com/ideas/archive/2022/10/against-algebra/671643/>
- Gutiérrez, R. (2018). Introduction: The need to rehumanize mathematics. In I. Goffney & R. Gutiérrez (Eds.), *Rehumanizing mathematics for Black, Indigenous, and Latinx Students* (pp. 1–10). NCTM. <https://math.oregonstate.edu/sites/math.oregonstate.edu/files/2022-03/Rochelle%20Gutierrez%20-%20The%20Need%20to%20Rehumanize%20Mathematics.pdf>
- IBIS World (2023). *Mississippi Economic Profile*. <https://www.ibisworld.com/united-states/economic-profiles/mississippi/#:~:text=What%20is%20Mississippi's%20GDP%20by,combined%2042.4%25%20of%20state%20GDP.>
- Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The journal of the learning sciences*, 13(2), 129-164. https://doi.org/10.1207/s15327809jls1302_1
- Kimbrough, K. (2021). Restaurant or your kitchen: Soup is good anyplace you find it. *Eat. drink. Mississippi*. <https://www.eatdrinkmississippi.com/post/soup-is-good-anyplace-you-find-it>
- Kroger (2023). <https://www.kroger.com/>
- Ladson-Billings, G. (2022). *The dreamkeepers: Successful teachers of African American children*. John Wiley & Sons.
- Marshall, S. (2023). But What Does it Look Like in Maths? A Framework for Culturally Sustaining Pedagogy in Mathematics. *International Journal of Multicultural Education*, 25(1), 1-29. <https://doi.org/10.18251/ijme.v25i1.3251>
- MDAC (2023). *Mississippi Agriculture Snapshot*. <https://www.mdac.ms.gov/agency-info/mississippi-agriculture-snapshot/>
- MSDH (2022). *Mississippi's Most Popular Names*. <https://msdh.ms.gov/msdhsite/handlers/printcontent.cfm?ContentID=25193&ThisPageURL=http%3A%2F%2Fmsdh%2Ems%2Egov%2Fmsdhsite%2Findex%2Ecfm%2Findex%2Ecfm&EntryCode=25193&GroupID=31#:~:text=Baby%20Girl%20>

- Names%20%C2%B7%20Ava%20142%20%C2%B7,Charlotte%2072%20%C2%B7%20Brooklyn%20/%20Emma%2068
- Meyer, D. (2010, March). *Math class needs a makeover* [Video file]. TED Conferences. https://www.ted.com/talks/dan_meyer_math_curriculum_makeover
- Meyer, D. (2015). Missing the Promise Of Mathematical Modeling. *The Mathematics Teacher*, 108(8), 578-583. <https://doi.org/10.5951/mathteacher.108.8.0578>
- Meyer, D. (2023). Three tests AI needs to pass before it can start to transform teaching. *Mathworlds*. <https://danmeyer.substack.com/p/three-tests-ai-needs-to-pass-before>
- NAEP (2022). *NAEP report card: Mathematics*. <https://www.nationsreportcard.gov/mathematics/sample-questions/?grade=8>
- NCTM (2020). Standards for the preparation of secondary mathematics teachers. <https://www.nctm.org/Standards-and-Positions/CAEP-Standards/>
- NCTM (2024). *Artificial intelligence and mathematics teaching: A Position of the National Council of Teachers of Mathematics*. <https://www.nctm.org/standards-and-positions/Position-Statements/Artificial-Intelligence-and-Mathematics-Teaching/>
- Pope, C. (2020). Ireland's favourite dinners revealed – and they're not quite what you'd expect. *The Irish Times*. <https://www.irishtimes.com/life-and-style/food-and-drink/ireland-s-favourite-dinner-s-revealed-and-they-re-not-quite-what-you-d-expect-1.4162531#:~:text=Invaluable%20book%20for%20the%20aficionado,to%20be%20entirely%20plant%2Dbased>
- Sapkota, B., & Bondurant, L. (2024). Assessing concepts, procedures, and cognitive demand of ChatGPT-generated mathematical tasks. *International Journal of Technology in Education*, 7(2), 218-238. <https://doi.org/10.46328/ijte.677>
- Staffel, J. (2023). AI in academia: Teaching challenges and opportunities. NCFDD. <https://www.ncfdd.org/januarynews24>
- TNTP. (2018). *The opportunity myth: What students can show us about how school is letting them down—and how to fix it*. https://tntp.org/assets/documents/TNTP_The-Opportunity-Myth_Web.pdf
- Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, opportunities, and implications for teacher education. *Contemporary Issues in Technology and Teacher Education*, 23(1), 1-23. <https://www.learntechlib.org/primary/p/222408/>
- Yu, H. (2023). Reflection on whether Chat GPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, 14, <https://doi.org/10.3389/fpsyg.2023.1181712>