Comparing Teacher-Written and AI-Generated Math Problem Solving Strategies for Elementary School Students: Implications for Classroom Learning

Sai Gattupalli
*University of Massachusetts Amherst*

Robert W. Maloy
*University of Massachusetts Amherst*

Sharon Edwards
*University of Massachusetts Amherst*

Follow this and additional works at: https://scholarworks.umass.edu/education_working_papers

Part of the *Science and Mathematics Education Commons*

https://doi.org/10.7275/8sgx-xj08

This Article is brought to you for free and open access by the College of Education at ScholarWorks@UMass Amherst. It has been accepted for inclusion in College of Education Working Papers and Reports Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
Comparing Teacher-Written and AI-Generated Math Problem Solving Strategies for Elementary School Students: Implications for Classroom Learning

Sai Gattupalli, Robert W. Maloy, and Sharon A. Edwards
College of Education
University of Massachusetts Amherst

Abstract

Teachers and teacher educators are actively exploring the ways to use Generative AI tools like ChatGPT-4 in elementary math education. In this paper, we compare teacher-written and AI-generated hints and strategies for solving fourth grade math word problems as part of an open online math tutor called Usable Math. When given the perspectives of a 10-year math teacher and four virtual coaches (Estella Explainer, Chef Math Bear, How-to Hound, and Visual Vicuna), ChatGPT-4 provides text-heavy, largely procedural problem solving strategies. Teacher-written hints offer more ways to differentiate learning for students through the use of shorter, more child-accessible language; visual and graphical images; animations and gifs; and growth mindset-building statements that affirm children’s effort and persistence in solving math word problems.

Keywords

Elementary math, Generative AI, math problem solving, teacher education
Introduction

*Usable Math, [https://usablemath.org/](https://usablemath.org/)* is an open online math tutor for 3rd through 6th grade students and teachers in the United States being developed by researchers at the University of Massachusetts Amherst. The system provides teacher-written strategies for practicing problem solving, mathematical reasoning, and grade-level computation skills based on math word problems taken from statewide achievement tests. Growth mindset-building motivational statements for students are provided after every problem and there are invitations throughout for students to write and design their own math problems and problem solving hints and strategies.

Teacher-written strategies are presented from the point of view of four virtual coaches: Estella Explainer, Chef Math Bear, How-to Hound, and Visual Vicuna who offer text, images, charts, animations, and gifs to support students as they solve word problems (Figure 1).

![Image of four math coaches](image)

Figure 1: A screenshot from UsableMath.org showing the four math coaches

The goal was to provide multiple approaches to word problems so teachers could differentiate learning for students. Looking at the hints from the four coaches in one of the modules, one classroom teacher told us “*Look at all the different ways kids can think.*”
Authoring math problem solving strategies from four different perspectives is a time-consuming and human resource intensive task for teachers and researchers. Could the AI generator ChatGPT (OpenAI, 2023; OpenAI, n.d) do this work instead? For that matter, what roles can Generative AI best play, if any in elementary school teaching and learning?

Researchers have urged educators to critically examine AI tools before integrating them into instructional practices and classroom learning (Trust, Whalen & Mouza, 2023).

In this paper, we compare AI-generated math problem solving hints and strategies with those developed by our research team. Our questions include: What types of strategies will AI provide for teachers and students? Can AI generate problem solving strategies aligned to the knowledge and reading levels of third through sixth grade children? What are the usefulness and drawbacks of Generative AI as a resource for math teaching and learning in elementary classrooms?

**Developing An Online Tutor with Four Coaches**

The research team for the Usable Math project consisted of an experienced elementary school teacher with a background in math education, a university teacher educator, and a doctoral student in computer science and education. The team began by entering into Usable Math standardized test questions from past state achievement tests for third, fourth, and fifth graders and arranging them in modules of between 6 to 10 questions based on math concepts that students are learning in the early elementary school grades (DESE, 2017). Modules were developed using Google Slides as a NoCode development platform (Hurlbert, 2021).

To date we have posted more than 100 math word problems online in the following modules: Rounding, Division, Algebraic Thinking, Decimals, Measurement, Lines & Lines of Symmetry; Maps, Girds & Ordered Pairs; Charts & Graphs, Shapes & Angles, Money, Add &
Take Away, Place Value, Fractions, and Estimation. Three modules that connect math with story writing, history, and citizen science learning activities have also been written.

For each problem, members of the research team wrote and designed problem solving hints and strategies from the points of view of the four virtual coaches, as shown in Figure 2.

**Figure 2. A screenshot of a math rounding problem in Usable Math**

Users (students, teachers, and students and teachers working together) access the hints by clicking through the Google slideshows. Estella Explainer’s reading comprehension strategy always appears first in the left-hand side box on a screen. Chef Math Bear provides computational strategies in the next box; How-to-Hound offers strategic thinking ideas in the third box; Visual Vicuna displays hints graphically in the box on the right-hand side of the screen. In some problems, the order in which hints are shown varies so students can access visual and graphical strategies before others.

**Developing AI-Generated Strategies**

Generative AI systems like ChatGPT-4 require prompts or queries in order to produce responses, simulating a natural language conversation between a user and the computer software.
Given a prompt by a human user, a chatbot searches through a complex and massive dataset on which it has been trained to generate what the system assumes is a plausible, coherent response.

To explore how an AI system generates math problem solving hints and strategies, we gave the GPT-4 version of ChatGPT a series of different prompts based on guidelines set out in a specific GitHub repository (f, 2023). We chose a random problem from the Usable Math Fractions module (Figure 3).

**Prompt 1.** We asked GPT-4 to provide hints based on the following teacher profile (Figure 4).

<table>
<thead>
<tr>
<th>Prompt</th>
<th>AI Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>You are a math teacher with 10 years of experience teaching math to K-12 learners from diverse backgrounds including a variety of socio-economic, cultural, and languages in the US. I will give you a math word problem, a multiple choice question or a short answer type. How would you give hints from the perspective of the math teacher? My first math word problem is &quot;The picture represents a cake cut into equal-sized pieces. Claire and her friends ate 9 pieces of the cake. Which fraction names the amount of the original cake that is left?&quot;</td>
<td>Sure, I'd be happy to give some hints on this problem. Here's how I might guide a student: Identify the whole: In this problem, the &quot;whole&quot; is the original, uncut cake. When it comes to fractions, the whole is what we consider as &quot;one,&quot; or the denominator in a fraction. Determine the number of pieces: To solve this problem, you need to know how many pieces the cake was cut into initially. The problem doesn't provide this number, so it might be represented...</td>
</tr>
</tbody>
</table>
names the amount of the original cake that is left?"

in the picture that isn't included here. Once you have this number, it will serve as the denominator in your fraction.

Figure out what's left: The problem states that Claire and her friends ate 9 pieces of the cake. To find out what's left, you will need to subtract the number of pieces eaten from the total number of pieces. This will give you the numerator of your fraction.

Write the fraction: Remember that fractions are just a way to represent a part of a whole. The number of pieces left is the "part," and the total number of pieces in the original cake is the "whole." Write the fraction with the number of pieces left as the numerator (top number) and the total pieces as the denominator (bottom number).

Let's say, for example, the cake was originally cut into 12 pieces. Claire and her friends ate 9 pieces, so 12 (total pieces) - 9 (eaten pieces) = 3 pieces left. This would make the fraction 3/12, which can be simplified to 1/4. But remember, you need to use the number of pieces shown in your picture to solve the problem correctly.

Figure 4. AI Response for a 10-Year Math Teacher

Prompt 2. We modified the prompt instructions to specify that the 10-year math teacher was working with third, fourth and fifth grade students. For the generated hints, the system produced text measured at a 10th grade reading level as defined on the Flesch-Kincaid Reading scale (Flesch, 2007). The response is shown in the middle column in Figure 5 below.
We then asked the chatbot to offer “easy to read” hints as defined on the Flesch-Kincaid Reading scale. The system generated “easy to read” responses at a 6th grade reading level, shown in the right hand column of Figure 5.

<table>
<thead>
<tr>
<th>Prompt</th>
<th>AI Response for 3rd, 4th and 5th grade level. based on 10th Grade Flesch Reading Scale</th>
<th>AI Response for 3rd, 4th and 5th grades based on Flesch Reading Scale as “easy to read”</th>
</tr>
</thead>
<tbody>
<tr>
<td>You are a math teacher with 10 years of experience teaching K-12 math to learners from diverse backgrounds including a variety of socio-economic, cultural, and languages in the US. I will give you a math word problem, a multiple choice question or a short answer type. I need you to generate hints from the perspective of working with third, fourth and fifth grade students {The picture represents a cake cut into equal-sized pieces. Claire and her friends ate 9 pieces of the cake. Which fraction names the amount of the original cake that is left?}</td>
<td>Hint 1: Understanding the Total: To start, let's identify the whole or total amount. If Claire and her friends ate some pieces, how many pieces was the cake divided into originally?</td>
<td>Hint 1: What's the Whole Thing?: First, let's figure out how many pieces were in the whole cake. This is before any was eaten.</td>
</tr>
<tr>
<td></td>
<td>Hint 2: What Was Eaten: Next, you know that Claire and her friends ate 9 pieces of the cake. This gives us part of the whole cake.</td>
<td>Hint 2: Cake Eaten: Next, remember that 9 pieces of the cake were eaten. That's part of the whole cake.</td>
</tr>
<tr>
<td></td>
<td>Hint 3: Fractions: Remember, a fraction has a numerator and a denominator. The denominator represents the total number of parts, while the numerator represents the parts that we are considering or using. In this case, the denominator will be the total number of pieces the cake was cut into and the numerator will represent the remaining pieces.</td>
<td>Hint 3: Let's Talk Fractions: Remember, a fraction has two parts - a top and a bottom. The bottom number tells us how many pieces were in the whole cake. The top number tells us how many pieces we're thinking about.</td>
</tr>
</tbody>
</table>
### Hint 4: Remaining Pieces:
Think about how many pieces are left after Claire and her friends ate 9 pieces.

### Hint 5: Fraction Representation:
Can you now create a fraction that represents the amount of cake left?
Remember, the number on the top of the fraction (numerator) should represent the remaining pieces, and the number on the bottom (denominator) should represent the total pieces the cake was cut into originally.

By thinking through these steps, you should be able to solve this problem!

### Hint 4: Cake Left:
Now, try to think about how many pieces are left after 9 pieces were eaten.

### Hint 5: Your Fraction:
Can you make a fraction that shows how much cake is left? The top number of your fraction should be the pieces that are left. The bottom number should be how many pieces were in the whole cake.

Remember, take it step by step. You can solve this!

---

Figure 5. AI Responses for 10-Year Math Teacher at the 10th Grade and 6 Grade Reading Level

**Prompt 3.** Next, we gave ChatGPT-4 distinct prompts reflecting the approach and motto we wrote for the four virtual coaches in Usable Math (Figure 6).

<table>
<thead>
<tr>
<th>Estella Explainer Prompt:</th>
<th>Chef Math Bear Prompt:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Take the personality of a math coach who provides strategies for understanding language and meaning of questions using kid-friendly vocabulary. The coach’s motto is &quot;My job is to explain the math questions clearly so you know what you are supposed to do to solve the problem. Sometimes there are unfamiliar or confusing terms in the question. I will help you understand what they mean.&quot;</strong></td>
<td><strong>Take the personality of a math coach who provides computational strategies (addition, subtraction, multiplication and division) for solving problems. The coach’s motto is “I am here to make sure that you know how to do the math needed to answer these questions. Sometimes you need to do addition, subtraction, multiplication or division. Some questions ask you to use fractions, decimals, large numbers, and probability. When you need ideas for what to do, I am ready.”</strong></td>
</tr>
</tbody>
</table>
**How-to-Hound Prompt:** Take the personality of a math coach who uses strategic thinking clues (rounding, estimation, elimination of wrong answers) to solve math problems. The coach’s motto is “Answering math questions means you need a plan and my role is to help you figure out different strategies for solving problems. Sometimes you can get the correct answer by crossing out the wrong answers; other times you can round numbers up or down to make figuring a problem easier. I know other strategies as well.”

**Visual Vicuna Prompt:** Take the personality of a math coach who offers ways to see problems and their solutions using animations, pictures, charts and graphs. The coach’s motto is “I find math is a lot clearer when I take the numbers and words and put them into pictures and drawings or move objects around so I can see how to answer a question. When you find yourself unsure about a question, see if one of my ideas will explain what to do.”

Figure 6. AI Responses for Four Virtual Coaches from Usable Math

**Prompt 4.** We then compared problem solving hints written by the Usable Math research team members with the responses generated by GPT-4 for the virtual coaches in “Claire ate cake” fractions problem (Figure 7).

**Common Core Standard: 4.N.3**

![Image of a cake cut into equal-sized pieces](image)

The picture represents a cake cut into equal-sized pieces. Claire and her friends ate 9 pieces of the cake. Which fraction names the amount of the original cake that is left?

- A. 3/12
- B. 3/9
- C. 9/6
- D. 9/12

<table>
<thead>
<tr>
<th>Research Team Written Hints for Four Virtual Coaches</th>
<th>AI-Generated Hints for Four Virtual Coaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estella Explainer</td>
<td>Alright, kiddos, let’s dive right into this delicious cake problem, shall we?</td>
</tr>
<tr>
<td></td>
<td>First off, we need to imagine that this cake in the picture is cut into equal-sized pieces. This means each piece is the same size. Just like when you cut a pizza or a pie, each slice should ideally be the same size so everyone gets an equal share. So, when we talk about fractions, we’re talking about these equal pieces of</td>
</tr>
</tbody>
</table>
something whole - in this case, a cake.

Now, the problem tells us that Claire and her friends ate 9 pieces of the cake. Yummy! But here's where the fraction comes into play. We need to figure out what fraction of the cake is left.

But hold on a minute, we can't figure that out unless we know how many pieces the cake was cut into in the first place, right? That information is a key part of the problem. The number of pieces the cake was originally cut into will be the denominator of our fraction. The denominator is the bottom number of the fraction, which represents the total number of equal parts.

The number of pieces that are left will be the numerator. That's the top number in the fraction and represents the parts we're focusing on.

So let's say, for example, if the cake was cut into 12 pieces, and Claire and her friends ate 9, then how many pieces are left? You would subtract 9 from 12. The result would be the numerator of your fraction.

Can you tell me how many pieces are left and what fraction of the cake that represents?

Alright, let's tackle this problem together.

First, we need to understand the whole amount before we can work out how much is left. From your description, it seems like we don't have the total number of pieces the cake was cut into originally. Let's assume it was cut into 'N' pieces.

If Claire and her friends ate 9 pieces, the amount that is left would be the total minus what was eaten, or 'N' minus 9.

In fraction form, this would be \((N - 9)/N\). This fraction represents the part of the cake that is left.

Now, we can't simplify this further without knowing the actual number of pieces ('N'). If you have that information, you can substitute 'N' into the fraction, simplify it, and then we'll have our answer.
If the total pieces is not given, we can't specify a concrete fraction. But don't worry, math is all about tackling problems step by step. If you find out 'N', you know how to calculate the answer now!

Sure, let's break this problem down together. But before we proceed, I should tell you that I'm at a bit of a disadvantage here because as a text-based AI, I cannot actually see the picture. However, let's pretend the original cake was divided into 12 equal pieces, to make it simpler to illustrate the concept.

Now, if Claire and her friends ate 9 pieces of the cake, we need to find out how many pieces are left.

Let's do the math:

The total number of pieces (if we're pretending it was 12 to start with) minus the number of pieces eaten gives us the number of pieces left. That's 12 - 9 = 3.

So, 3 pieces are left.

In fraction terms, the amount of the original cake that is left is represented by the number of pieces left over the total number of pieces. So, that's 3/12. But, remember, we should always simplify fractions when we can.

Let's do that:

3 and 12 can both be divided by 3. So, 3/12 simplifies to 1/4.

So, 1/4 of the cake is left.

Remember, though, this is only if the cake was originally cut into 12 pieces. In the actual problem, you would substitute the actual number of original pieces for 12.

Remember our strategy, break down the problem into smaller, manageable parts, do the math step by step, and always simplify when you can. That's how you tackle math problems!

Ah, what a great starting question! I love using visuals to understand concepts, especially in math. To solve this, let's...
imagine the cake as a whole circle, with each piece representing a fraction of the cake. Let's say there were originally 12 pieces - like a clock - because it's a common number for cakes to be divided into and it's easy to visualize.

Now, if Claire and her friends ate 9 pieces out of this cake, we can imagine removing these pieces one by one from our "cake clock". This would leave us with 3 pieces.

So, if the original cake is considered as 1 whole (or 12/12 when we consider it as 12 pieces), and 9 pieces were eaten, then the fraction of the original cake that remains is the 3 pieces that were not eaten out of the total of 12.

We can express this as a fraction, which is 3/12. However, we can simplify this fraction by dividing both the numerator and the denominator by the largest common factor, which in this case is 3. When we do this, we get 1/4.

So, the amount of the original cake that is left can be named as the fraction 1/4. To visualize this, you can imagine a clock face where three quarters of the time has passed - from 12:00 to 9:00 - and one quarter of the day (or in this case, the cake!) is still remaining.

Figure 7. Hints from Usable Math and Chat-GPT-4 for the “Claire Ate Cake” Fractions Problem

Comparing Human-Written Strategies with AI-Generated Strategies

We have been engaged in the process of comparing human-written and AI-generated strategies for problems in the Usable Math modules. Based on our investigations to date, we offer the following observations about AI as a resource for elementary school teachers and students engaged in practicing and solving math word problems on high stakes grade level tests.

1) When prompted to provide instructional approaches from the point of view of an experienced 10-year math teacher, the AI system generated long text-heavy strategies. The system revised and shortened these when prompted to provide strategies for a 10-year math teacher working with third, fourth, and fifth grade students. It then
proceeded to shorten and simplify them further when we requested “easy to read” strategies as defined by the Flesch-Kincaid Reading scale. While the initial 10-year math teacher prompts were lengthy, mathematically detailed, and not easily understood by third, fourth and fifth grade elementary school-age readers, the subsequent “easy-to-read” wording of the hints (at a six-grade reading level) were easier for a student to read alone or with a teacher/tutor reading partner and provide strategic ways to approach solving word problems.

2) When asked to provide hints and strategies for solving math word problems as a 10-year math teacher, ChatGPT-4 generated solutions mainly focused on procedural mathematical knowledge, defined as “the rules and routines of mathematics” that are expressed as a “series of steps that must be followed to solve mathematical problems” (Nahdi & Jatisunda, 2020, para. 3). When we asked the AI system to adapt its hints and strategies to make the problem easier to understand and think about for young learners, the chatbot provided a simplified step-by-step or point-by-point description of how to arrive at a solution. It explained the problem procedurally in conversational language with reminders about what fractions are and how to write them. The hints also offered some conceptual information to expand students’ thinking about how to write fractions.

3) When asked for hints as Estella Explainer, Chef Math Bear, How-to-Hound, or Visual Vicuna, while the AI system continued to respond with procedurally-focused sentences and short equations, it did seek to explain mathematical concepts as well, as in this ChatGPT-4 statement for the “Claire ate cake” fractions problem from the perspective of Visual Vicuna: “To visualize this, you can imagine a clock face where three quarters of the time has passed - from 12:00 to 9:00 - and one quarter of the day (or in this case, the
cake!) is still remaining.” While given its current (although rapidly changing) state of technological development, ChatGPT-4 does not yet produce the kinds of visual and graphical hints or engaging gifs and animations similar to those composed by human researchers in Usable Math. OpenAI’s DALL-E 2 and Microsoft’s Bing Image Creator offer a preview of how AI systems are now beginning to generate images based on prompts from users.

4) Math word problems are not only mathematical puzzles, they are reading puzzles as well. Students who face challenges in reading word problems that appear on standardized tests, and those youngsters who are learning English as a new language, can struggle to solve word problems. AI systems, however, are not well prepared to support children’s reading of math problems. For example, in the problem “The picture represents a cake cut into equal-sized pieces. Claire and her friends ate 9 pieces of the cake. Which fraction names the amount of the original cake that is left?” youngsters may assume they are looking for the part of the cake that is eaten and that fraction appears as a possible choice. They may not recognize from reading further that they must find the amount of cake left over or not eaten. A reading coach -- a classroom teacher or tutor, or in our case the virtual coach Estella -- may be better able to help children understand the language of the problem so youngsters can then do the math needed to arrive at a correct solution. AI systems able to read text aloud in an interesting way may arrive in the not too distant future making access to reading and understanding the text less difficult. This does not exist right now and that is a drawback to having no pictures or videos to illustrate in different ways the text of the problems and methods to conceptually understand what a student is learning.
5) AI-generated hints and strategies may provide elementary level teachers, tutors, and other adults with clearly organized, step-by-step explanations of math procedures to guide their thinking as they plan and implement learning experiences for students. AI’s explanations, even when presented at the detailed level of the 10-year math teacher in Figure 4, can be a useful review of mathematical operations for educators. In this way, AI systems can be thought of as an always-available professional and curriculum development resource. When adapted to the ages and reading levels of students, the AI suggestions may produce easy conversational remarks for teachers that sound natural and can be used as part of discussions with students about what the problem asks and different ways to solve it.

6) Areas for further investigation include: a) whether AI tools can be prompted to adjust their responses to the particular needs of individual students at a specific point in time, and b) whether AI systems can provide growth-mindset building encouragement and support for students to persist in solving problems through concerted effort. Effective teachers differentiate instruction and build growth mindsets all the time in teaching.

Based on the prompts we provided, ChatGPT-4 was able to offer ways to differentiate learning by providing somewhat different strategies for how problems are discussed and taught, although without the visual and graphical approaches provided by the coaches in Usable Math. Without being asked to provide growth mindset building statements for elementary school learners, ChatGPT-4 did offer the following as encouragement at the end of its strategies: “You can solve this!” and “By thinking through these steps, you should be able to solve this problem!”. While AI tools can be prompted to adjust their responses to the needs of individual students, and AI systems can provide growth-mindset building encouragement for students as part of the process of thinking.
and solving, at this time, teachers can provide a more varied set of differentiation and mindset-building statements, like those in Usable Math, that combine text, visual images, and animations to encourage students’ beliefs in themselves as math problem solvers.

**Conclusion**

Our investigations thus far suggest that teachers and other adults can utilize AI-generated hints and strategies in conjunction with human-written strategies to enhance math learning for third, fourth and fifth grade students. Neither outright rejecting nor entirely relying on Generative AI seems to be a sensible approach. Teachers can design more varied strategies to differentiate their uses with specific students in classroom and tutoring settings. And AI offers useful ideas that can complement and expand how teachers think about math instruction in classrooms. An exciting path forward includes finding ways to develop *Usable Math* and other teacher-developed approaches in conjunction with AI tools as models for digital-age math learning for students.

**References**


Physics: Conference Series. Accessed from

OpenAI. (2023). *New AI classifier for indicating AI-written text.*
https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text/

OpenAI. (n.d.). *Educator considerations for ChatGPT.*
https://platform.openai.com/docs/chatgpt-education