

Math Isn't Neutral: Designing Word Problems with GPT-4 for Relevance*

Mosum Trivedi

Wayne State University

Abstract

Textbook word problems often miss students' lives; AI can fix that only when teachers stay in the driver's seat. This article shows a practical, repeatable way to use GPT-4 to design mathematically rigorous tasks that feel relevant to students. I present an iterative prompting approach that pairs content goals with two added parameters: [SC] Social Context and [PA] Pedagogical Approach, alongside the familiar task elements (object, shape, properties, target, position). Drawing on examples from a Detroit high school, the paper traces how simple geometry items were reshaped into modeling tasks, civic case studies, and a systems-based investigation of environmental data. For immediate classroom use, the article includes: (1) a step-by-step "try it next week" recipe, (2) a revision checklist for catching quantity/wording issues, and (3) a rubric for analyzing the results. A brief classroom pilot (a trigonometry quiz with local contexts) illustrates gains in student talk, diagramming, and flexible reasoning, as well as common pitfalls and how to revise AI drafts. The goal is not to automate curriculum, but to amplify teacher judgment and help answer the daily question, "Why are we doing this?", with tasks that connect mathematics to the world students inhabit.

Keywords: word problems, pedagogy, post-constructivism, relevant mathematics

1 Introduction

The development of artificial intelligence (AI) in education has sparked both excitement and concern, particularly within mathematics instruction. Tools like GPT-4 offer new possibilities for generating tasks, lesson materials, and scaffolds. But their usefulness depends not on novelty or automation, but on how thoughtfully they are integrated into teaching and learning. This paper explores the potential of AI as a collaborator—one that can assist educators in designing rich, meaningful, and contextually relevant learning experiences.

This paper investigates how GPT-4 can be prompted through iterative design to generate mathematics word problems that reflect our students, including their cultural, civic, and historical context. Traditional textbook problems are often written for a generalized, decontextualized audience, and as a result, frequently fail to connect with students, particularly in transferring into workable skills outside of the classroom.

The difficult task is to combine a modeling approach that contains meaningful and relevant context, with procedural practice that still allows for student choice. With deliberate prompting, GPT-4 can

*Generative AI tool ChatGPT was used for editing, organization, and brainstorming during the writing process. All AI-generated suggestions were critically reviewed, revised, and integrated by the author. The final manuscript reflects the author's original ideas, interpretations, and academic contributions.

help produce high-level mathematical problems that can engage students beyond rote procedure. Due to the differentiation needed for each classroom, that is, the different needs of different students, AI can be a thinking partner to create problems to overcome the issues of teacher-time restrictions and rigidity of traditional textbooks.

Building on this foundation, I use GPT-4 as a tool in a design-thinking iteration process to explore how it can help teachers modify existing curricular problems or generate new ones grounded in student agency and context but highly attuned to mathematical standards. Drawing on examples from a Detroit high school classroom, this paper outlines the process of developing and refining word problems that aim to blend mathematical rigor with contextual relevance. My intention is to inspire mathematics teachers to revisit pedagogy with a practical, repeatable way to design for relevance using AI.

2 Literature Review

Mathematics education has long struggled with a disconnect between the problems students are asked to solve and the worlds they inhabit. Word problems in textbooks are often decontextualized and generic, designed to be widely applicable and yet frequently irrelevant and even harmful (Buck & Damico, 2021). As Hwang and Utami (2024) note, “existing math word problem generation systems generally lack a ‘context,’ which refers to the real-world scenario in which the math equation is situated.” Despite calls for contextualization from researchers and professional mathematics organizations, “many schools continue to present mathematics as abstract, decontextualized problems” (Verschaffel et al., 2020).

To address this gap, educators have developed tools to evaluate whether a task supports relevance and equity in math learning. Aguirre and Zavala (2013), for example, propose a Culturally Responsive Mathematics Teaching (CRMT) framework that includes five key dimensions for task analysis: Mathematical Depth, Mathematical Discourse, Student Engagement, Contextual Connection, and Critical Insight.

These dimensions provide a structure for teachers to evaluate curriculum, but they also offer a useful lens for examining whether emerging technologies like AI can generate tasks that are not only mathematically sound, but socially meaningful.

Generative AI systems like GPT-4 can produce mathematics word problems almost instantly, and adjust for difficulty and structure. Hwang and Utami (2024) investigated this capacity using a prompting framework that helped GPT generate geometry problems across three levels of complexity: easy, intermediate, and advanced, based on the number of parameters included in the prompt. Their parameter options are as follows: [O] object detected, [S] shape recognized, [P] measured properties, [TP] target of the problem, and [POS] position for compound shape. Their study confirmed that GPT-generated problems could align with content standards and difficulty expectations, and this paper carries on the next step, to examine the quality of the problems’ contexts and their cultural or social resonance. Without this examination, generative AI may replicate the same faux-neutral, disengaged framing as traditional curriculum unless educators intervene. Prompting alone is insufficient if the prompt is not grounded in pedagogical purpose and social context.

While Aguirre and Zavala’s CRMT framework offers a powerful tool for evaluating classroom tasks, it is not the only pathway for making mathematics relevant. Abah (2017) offers a complementary approach, arguing that incorporating mathematical history into curriculum design can help students see math as a social endeavor, emergent from the development of human society, and continuing to further its progress. He critiques ahistorical teaching for reinforcing alienation and shows how even foundational mathematical procedures can be enriched by historical grounding. This approach avoids

relying solely on ethnic or identity-based relevance, offering instead a connection to the broader arc of human development and problem solving.

In addition to frameworks centered on cultural relevance, Lesh et al. (2003) advance a post-constructivist ‘models and modeling’ perspective that shifts the focus of mathematics instruction from solving to constructing. They emphasize that students learn meaningfully when they create, test, and refine models to represent complex systems using multiple forms of representation. This approach complements culturally grounded frameworks by insisting that problems must also support mathematical reasoning and decision-making. When paired with tools like GPT-4, modeling tasks offer a pathway for designing problems that centers not just on relevance, but on student agency and epistemic engagement.

Taken together, these foundations guide how to prompt and evaluate AI-generated problems that move beyond mere standards alignment to support contextual relevance, inquiry, and engagement. Their combined influence also surfaces a key challenge: designing prompts that embed both the students’ immediate experiences and the larger historical and social forces shaping their lives.

3 Design Approach

This paper employs post-constructivism and design thinking as both pedagogical lenses and design processes to foreground student relevance and agency.

Building on these frameworks, my contribution involves developing an evaluation lens that combines social relevance, historical context, and civic knowledge. I adapt Aguirre and Zavala’s original dimensions to include a broader conception of relevance: one not limited to identity, but grounded in student engagement and reality.

- **Cognitive Depth:** Does the task require reasoning, modeling, or conceptual understanding?
- **Mathematical Discourse & Flexibility:** Does it invite explanation, justification, and multiple solution paths?
- **Student Engagement & Ownership:** Does it create space for curiosity, creativity, and personal choice?
- **Contextual Relevance:** Does the problem reflect timely or familiar references (e.g., local history, civic issues, or pop culture)?
- **Insight into Society:** Does it prompt reflection on social reality?

What follows in Section 3 is an iterative series of GPT-4 prompts, showing the process of developing GPT-4 to create problems in support of this broader framework of relevance. Beginning with standard geometry problems and expanding to include Detroit-specific cultural references, historical landmarks, and civic controversies, I evaluate each task using the revised dimensions above. Rather than focusing solely on cultural identity, the goal is to emphasize math as a human, historical, and applied tool that students can use to understand and influence the world around them.

To build on the work of Hwang and Utami (2024), I add a new parameter:

[SC]: Social Context

This adaptation shifted the focus from mathematical accuracy alone to the possibility of relevance-driven problem generation. My goal was not only to generate scaffolded math tasks, but to explore whether GPT-4 could create problems situated in meaningful local, historical, political, or conceptual contexts, which could activate curiosity, inspire dialogue, and spark insight. However, to overcome the ‘window dressing’ issue, the pedagogical approach must also be input by the teacher. In this case, design learning and modeling will be the style used to support students in using mathematics

to interrogate and explain real conditions. This requires the teacher's own pedagogical approach to become explicitly named in the prompt. I thus add a new parameter:

[PA]: Pedagogical Approach

This parameter can be changed, thus allowing teachers to become an active designer of the classroom, rather than a passive promoter of the decisions of textbook writers.

4 Iterative Design Results

I encourage readers to open a GPT window and follow along with the prompts, changing them to your particular contexts, including location and mathematical subject.

We will use different combinations of the following parameters in order to review the output:

Framework Component	Description / Function
[O] Object	Concrete element or item anchoring the problem
[S] Shape	Geometric figure identified
[P] Properties	Measured or inferred quantities
[TP] Target	What students must solve or calculate
[POS] Position	Spatial or relational orientation
[SC] Social Context	Place, history, politics, or social condition
[PA] Pedagogical Approach	Modeling, Design Learning, Inquiry, etc.
Relevance Dimensions	Cognitive Depth, Mathematical Discourse & Flexibility, Engagement, Contextual Relevance, Insight into Society

Prompt: "I will use this chat to create geometry word problems. I will input different parameters in order to control the difficulty level, and expect the word problem and the answers."

To further support the aims listed above, I asked GPT to retrieve and summarize background information on Detroit's history, demographics, political developments, and cultural landmarks.

Prompt: "Tell me about life in Detroit in 2024 including the history, population demographics, government plans, and culture"

I also included another prompt in the interest of expanding student identity beyond immediate geography.

Prompt: "Add world history, American history, and mathematical history as further options for contexts for making geometry word problems".

This additional context helped shape subsequent prompts, allowing for deeper and more responsive design. The process was intentionally iterative: after each generated problem, I reflected on its strengths and limitations across the revised five-part framework and refined the prompt accordingly.

While the examples I show here use Detroit as an anchor, the methodology applies across contexts. The goal is not parochialism, but to situate mathematical modeling within systems, whether local, state, national, or global, that reveal and reflect on processes that affect students.

4.1 From Standard to Local: Initial Prompt

Parameters: Shape, Target, Properties, Social Context

Prompt: "Create a geometry word problem about the area and perimeter of a rectangle with length 150 cm and width 100 cm, with relevance for students that live in Detroit in 2024."

GPT-4 Output:

The City of Detroit is planning to create a new community garden in a rectangular plot of land in a neighborhood park. The garden will have a length of 150 cm and a width of 100 cm on the scaled blueprint (where 1 cm = 1 meter in real life).

1. What will be the actual area of the garden in square meters?
2. What will be the actual perimeter of the garden in meters?
3. If the city wants to place a fence around the garden, costing \\$25 per meter, how much will it cost to fence the entire garden?

Reflection: This prompt was a baseline, revealing the limitations of default GPT responses, which use tacked on, surface-level context. The forcing of physical parameters also creates a closed-ended problem rather than an open-ended problem where students must take ownership of the process, which certainly has its place in procedural practice, but can be found in traditional textbooks and does not need AI to create. Due to a lack of pedagogical instructions, the social context is performative rather than a means of creating student buy-in.

4.2 Anchoring in Landmark and Symbolism

Prompt: “Create a geometry word problem about a stacked compound shape of a rectangle and a trapezoid with cultural relevance to Detroit students in 2024.”

Parameters: Shape, Position, Social Context

GPT-4 Output:

The Charles H. Wright Museum of African American History in Detroit is planning a new sculpture for its outdoor plaza. The sculpture’s base will be a rectangular section representing the museum’s solid foundation of preserving African American history. On top of the rectangle will sit a trapezoidal section symbolizing the diversity and unity of Detroit’s cultural legacy.

The rectangular base has a width of 15 meters and a height of 20 meters. The trapezoidal section has a height of 10 meters, with the shorter base measuring 8 meters and the longer base measuring 15 meters.

1. Find the total area of the sculpture’s front face, combining the rectangle and the trapezoid.
2. If the front face is to be covered with a reflective gold material costing \\$6 per square meter, what will be the total cost?
3. The museum plans to install decorative LED lights around the perimeter of the sculpture’s front face. Calculate the total length of the lights needed.

Reflection: This prompt attempted to peel back some of the parameters to see if it would improve the output. This response shows a slightly improved grounding in community context and symbols. It is still teacher directed, and introduces some narrative elements and asks students to engage in a multi-directional manner given different parameters. Students are asked to create some models, but there will be exactly one correct answer, revealing a lack of student ownership. Social context is still not embedded in any meaningful way.

The next iterative prompt remained closed-ended but embedded rich civic context and discussion.

GPT-4 designed problems like a mural budgeting scenario that invited students to reflect on public art funding. While not fully open-ended, this type of task can still support contextual insight and spark cross-disciplinary thinking.

Prompt: “Create a geometry word problem involving area and perimeter that includes an open-ended civic discussion prompt and references public revitalization efforts in Detroit.”

Parameters: Target, Social Context, Pedagogical Approach

4.3 Civic Modeling with Trigonometry

Parameters: Pedagogical Approach, Target, Social Context

Prompt: “Create an open-ended problem that requires modeling where students must use trigonometric ratios and properties like special right triangles, and the Pythagorean theorem, with context of Detroit environmental issues, with an introductory anecdote at the beginning and discussion questions”

GPT-4 Output:

Task: Monitoring Air Quality in Detroit

In recent years, Detroit has made progress in revitalization, but environmental concerns still affect many neighborhoods—especially those near industrial zones. Air pollution, particularly from factories, highways, and freight routes, can be harmful to residents’ health. Local scientists and community activists are working together to track and model how pollution spreads through neighborhoods to push for cleaner air policies.

One method environmental engineers use to measure pollution involves placing air quality sensors on buildings and calculating distances, heights, and angles between sensors and sources of pollution—often using special right triangles, trigonometric ratios, and the Pythagorean Theorem.

Today, you’ll take on the role of a community-based environmental engineer modeling pollution spread in Detroit. You’ll use trigonometry to design a monitoring strategy for measuring and visualizing air quality near a local highway or industrial site.

Open-Ended Modeling Task:

You are designing a network of air quality sensors to monitor emissions from a diesel-fueled freight hub near a school in Southwest Detroit. Your job is to determine:

- Where the sensors should be placed
- How high each sensor must be
- What angles they need to monitor effectively
- And how far pollution might drift under certain wind conditions

You’ll use this information to make a recommendation to the community board about whether pollution levels might exceed safe limits and where more sensors or changes might be needed.

Your Tasks:

1. Choose a Study Area:

Select a site in Detroit where there is environmental concern

- Sketch a rough map (can be abstracted to a coordinate plane).

- Choose three key locations for placing air quality sensors (on schools, homes, light poles, etc.).
- Identify at least one pollution source.

2. Use Trigonometry to Model Sensor Placement:

For each sensor, determine:

- The height of the sensor from the ground (using building or pole height).
- The angle of elevation or depression from the sensor to the pollution source (you may use estimated or provided distances).
- Use trigonometric ratios (sin, cos, tan) and special right triangles (e.g., 30° - 60° - 90° , 45° - 45° - 90°) to calculate: (1) Horizontal distances, (2) Heights of emissions, and (3) Line of sight ranges.

3. Use the Pythagorean Theorem to Estimate Drift:

Assume pollutants can drift diagonally with the wind. If pollutants rise 10 meters and drift horizontally 24 meters, calculate the actual distance they travel using the Pythagorean Theorem.

Repeat this with different wind directions and sensor locations.

4. Create a Model:

- Draw your site to scale on graph paper or using software.
- Label all distances, angles, and sensor placements.
- Include at least two triangles where you clearly show your trig and Pythagorean reasoning.
- Color code your diagram to show “high concern” and “low concern” zones based on how close they are to pollution sources and line-of-sight angles.

5. Recommend a Plan:

Write a one-paragraph recommendation to the city council or community board. Include:

- Where to place new sensors
- Whether pollution might reach the school or homes
- What mitigation strategies might help (e.g., tree barriers, traffic changes, sensor alarms)

Discussion Questions:

1. How did using trigonometric ratios help you make decisions in your model?
2. What did your model suggest about how far and high pollution can travel in Detroit’s neighborhoods?
3. Why is it important for communities to have access to accurate environmental data?
4. What are the limitations of your model? How could real-world data improve it?
5. What’s one action you think the local government or industry could take to reduce pollution exposure?

Reflection: This task reveals the GPT can create engaging projects that challenge students to create models, and in a context that matters to them. I maintained the ‘Detroit’ aspect for student buy-in but brought in a broader social issue. The problem blends spatial geometry and trigonometry with civic learning, offering students both voice and purpose. The project supports interdisciplinary thinking, fosters awareness of social issues, and models how mathematical tools can inform official policy. Finally, it reveals GPT’s strength to implement different pedagogical approaches, in this case modeling and relevance by placing the student as a driver of the choices. However, this task does maintain some closed-ended parts.

4.4 Expanding Student Ownership through Iteration

The design-learning, modeling approach that creates full student ownership was not met in the last iteration. Razzouk and Shute (2012) comment that “preparing students to succeed in the world, we should not require that they memorize facts and repeat them on demand; rather, we should provide them with opportunities to interact with content, think critically about it, and use it to create new information”. So, I iterated upon the last problem to ensure further student ownership:

Parameters: Pedagogical Approach, Target, Social Context

Prompt: “Continuing from the last problem, there are still closed problems that should be opened and have scaffolded instructions for students to follow.”

GPT-4 Output:

“Where Does the Air Go?”

Framing the Problem:

As Detroit continues its efforts to grow and redevelop, some communities remain more exposed than others to nearby highways, truck routes, and industrial activity. Air quality concerns—especially near sites like I-75, Zug Island, and freight corridors—have raised questions about how environmental conditions affect the health and safety of different neighborhoods.

Community organizations, scientists, and local residents are working together to better understand how pollutants move through city environments. Mathematical models can play an important role in helping interpret and visualize these patterns.

Your Task:

You are part of a youth-led data and design team working to model how air pollution might travel from a known source (such as a freeway or warehouse complex) to nearby homes, parks, or schools. You’ll use geometric modeling and trigonometry to build a visual representation of how pollutants might spread—and make recommendations based on your findings.

Project Steps:

Step 1: Frame the Situation

Choose a Detroit-based location where pollution might originate. Example sites:

- A freeway near a school
- A truck depot near homes
- A factory near a park

Then:

- Identify at least two locations nearby that might be affected.
- Assign or estimate coordinates for all points using a grid or scaled map.
- Include elevation, if relevant (e.g., second-story vs. ground level).

Reflect:

- What health or environmental concerns might people in this area have?
- What factors (wind direction, building height, topography) influence how pollution moves?

Step 2: Build Your Model

Create right triangle diagrams to represent how pollution might travel from the source to various points

in the community. You might:

- Use an estimated angle of elevation and a horizontal distance to calculate height
- Use a known height and an angle to estimate how far pollutants may travel
- Explore how triangle dimensions might change with different weather or wind conditions

Include:

- At least two labeled right triangles
- Written explanation of your triangle setup and calculations
- Application of trig ratios (sine, cosine, tangent) or the Pythagorean Theorem

Reflect:

- What do your calculations suggest about the exposure risk in the neighborhood you modeled?
- Are certain locations more protected than others?

Step 3: Evaluate and Adjust

Use mathematical reasoning to explore the sensitivity of your model:

- How would pollution travel under different wind angles?
- How far would it drift if the release height changed?
- If you had real data, where would you place sensors to test your model?

Reflect:

- What additional data would improve your model?
- What are its strengths and limitations?

Step 4: Present and Recommend

Create a short presentation or poster that includes:

- Your triangle-based pollution model
- Labels and measurements
- A brief explanation of your reasoning
- A recommendation for further data collection, design change, or community response

Optional: Make a recommendation for air sensor placement or other design choices based on your model.

Final Discussion Questions:

- What did trigonometry allow you to see or analyze more clearly?
- Why might it matter to model how pollutants move in a real place?
- How can math help communities plan, protect, or respond to environmental concerns?
- What questions do you still have about this issue?

Reflection: This iteration reveals how AI will respond to teacher prompts to create rich, interdisciplinary, and relevant modeling projects. Requiring AI to provide support for open-ended problems thus allows for student voice, positioning students as designers of the world rather than passive solvers. The teacher's evaluation of this project can then add or remove support, depending on the needs of the classroom. Some students may benefit from constraints, but if the classroom has been practicing modeling before this, then they may be able to manage the designing required for this open project.

4.5 Anchoring in Systems-Based Approach

Prompt: “Redesign this problem with a systems-based approach”

Parameters: Pedagogical Approach, Target, Social Context

GPT-4 Output:

Modeling Environmental Health Systems Near Industrial Corridors

Essential Question:

How can we model the interconnected systems that determine air quality in residential areas located near industrial trucking zones, and what actions can communities take to intervene?

Scenario:

In many U.S. cities—including Detroit, Houston, and Fresno—residential neighborhoods are located near trucking depots, highways, or industrial corridors. These locations increase exposure to PM2.5 (particulate matter), which contributes to asthma, heart disease, and other chronic illnesses. The amount of pollution a neighborhood experiences is not just about distance—it's about wind direction, traffic density, zoning laws, and disinvestment.

Your team is part of a community environmental health task force working to analyze and present spatial data about air pollution exposure. You'll be using geometry and trigonometry to model the distance and direction between neighborhoods and pollution sources, as well as estimating exposure zones. Then you will make recommendations for system-level changes.

Inputs and Parameters (Given):

- The truck depot is located at coordinates (0, 0).
- Neighborhood A is located 6 km northeast at a 45° angle.
- Neighborhood B is directly east, but closer to the depot.
- The prevailing wind direction is from the southwest, pushing pollution northeast.
- A new elementary school is being proposed in the area.
- The wind blows pollution up to 7 km in the wind's path.

Your Tasks (Student Instructions):

1. Model the Locations

Use a coordinate plane to model the truck depot, Neighborhoods A and B, and the proposed school.

2. Estimate Direct Distances

Use the Pythagorean Theorem and/or trigonometric ratios to calculate the distances from the depot to each neighborhood.

3. Model the Pollution Zone

Use trigonometry to sketch the pollution plume in the direction of the prevailing wind. Determine which neighborhoods and landmarks fall within that triangle.

4. System Mapping

Draw a diagram (or use a digital tool) to show the causal connections between pollution levels, zoning decisions, infrastructure investment, and public health outcomes. Use arrows, labels, and text to describe how parts of the system affect each other.

5. Policy Intervention Options

Choose one of these three interventions:

- Add more green infrastructure between the depot and neighborhoods.
- Relocate the depot further away.
- Create stricter air monitoring and community health alert systems.

6. Model Your Intervention

Model how your chosen intervention changes the geometry or math of the problem. (E.g., how far must the depot move to keep schools out of the pollution zone?)

Discussion Questions:

1. How do physical distance and angle of wind influence exposure risk?
2. Which variables in this system are controllable, and which are structural?
3. How does this model reveal systemic inequities in city planning?
4. What are the math-based limitations of this model—and what else would we need to make better policy decisions?

Reflection: This problem is not open-ended in the mathematical sense, but is open-ended in considering the application of specific proposals that reflect the civic proposal process. Prompting could change this to a fully open-ended problem, or keep it closed for students who are not ready for full-on modeling, or need scaffolding towards independent modeling. The framing and discussion questions can also become sharper in response to social issues, but that will be left as an exercise.

The exercises presented as examples do require the teacher to plan ahead with scaffolding, vocabulary, and modeling practice. AI can assist with the background work as well, including providing the vocabulary definitions, outlining misconceptions that might arise, and writing teacher planning guides. We can consider the broader design principles in developing the prompts that will bring forward these varying demands.

5 Design Principles for Contextual Word Problems Using GPT-4

For educators looking to explore contextual math problem creation with GPT-4, the following principles may support authentic experimentation:

- Start with both content standards and pedagogical intent
- Layer in civic, cultural, or historical questions as the context
- Include places for student decisions—modeling, design, justification, reflection
- Review each output for coherence in quantities, language, and internal logic
- Use AI-generated tasks as rough drafts: revise for tone, flexibility, and depth
- Let prompts evolve by considering what fundamental question is changeable

Pedagogical Approaches to test out:

<ul style="list-style-type: none"> • Mathematical Modeling • Direct Instruction • Design Learning • Project-Based Learning • Real-World Case Study 	<ul style="list-style-type: none"> • Cross-Curricular Integration • STEM/STEAM Integration • Socratic Dialogue • Historical Inquiry
---	---

Further ideas:

Prompt: Create the prompts for a Socratic Dialogue for area and circumference of a circle with relevance to students in Detroit

Prompt: What are ideas for Cross-Curricular Integration for Geometry topics?

Prompt: Create a project based on historical based inquiry regarding rigid transformations

6 Reflections on Prompting and Professional Judgment

This prompting process made one thing very clear: we cannot use basic prompts and expect an improvement over existing materials. The early prompts imitated textbook word problems: surface-level “real-world” settings that changed the nouns without changing the role of the student in the problem.

The full modeling projects in 3.4 and 3.5 pushed toward what I wanted: open-ended, locally grounded, and required students to reason through their own choices. For students who haven’t practiced modeling before, those would be a jump. So I prompted GPT to scaffold backwards:

Prompt: “Set up two mini-projects that will help scaffold up to the larger modeling project on environmental pollution.”

The results were useful for draft material for shorter, more focused modeling activities that help students build fluency with tools like trig ratios and triangle setup before we dive into a full project. This is the purpose: AI can accelerate brainstorming and help build raw material, and even draw teachers into avenues not previously considered, and create the type of engagement that adult practitioners of mathematics will use.

7 Informal Classroom Implementation: Trigonometry Quiz with Social Context

I have been using GPT-4’s utility in generating different tasks with high engagement and relevance. In one use, I prompted for different versions of a trigonometry quiz rooted in Detroit specific contexts. AI featured references to local figures such as Detroit Pistons players and fashion designers, as well as recognizable landmarks as relevant context where the trigonometry was found. While this quiz was intended to check procedural fluency rather than modeling or inquiry, the questions still encouraged visual reasoning and flexibility.

GPT’s initial outputs for the trigonometry quiz problems were lively and varied, but I prompted for parameters to shift, and GPT-4’s writing was confusing for a student, using unclear subject references. I revised the questions to maintain consistent quantities while preserving the structure of varying the opposite, adjacent, or angle to ask: “What happens if this changes?” and “What is the maximum or minimum this could be?” This required careful oversight, but the base ideas provided a creative springboard.

In the quiz’s implementation, students were grouped into fours, and in the groups, were highly engaged in discussion about planning, execution, and results. Many spent more time drawing triangle models than they had in previous assessments, and several commented on how the changing variables made them think differently about trigonometric relationships. The quiz successfully combined structured fluency practice with relevant context and conceptual flexibility, demonstrating that even closed-ended GPT-generated tasks can offer learning opportunities when consciously curated.

8 Implications for Teaching and Learning

The work presented in this paper suggests that AI tools like GPT-4 can do far more than accelerate planning or creating simple warm-ups, by generating math activities that reflect the world. However, realizing this potential requires a shift in mindset: from using AI as a tool of automation to a tool of intellectual and pedagogical engagement.

GPT-4 can propose endless numbers of tasks, but teachers must:

- Filter them for mathematical coherence
- Adapt them for their students' interests and contexts
- Embed them into intentional pedagogical structure

The teacher's role in reviewing output cannot be overstated. When AI creates a problem it does not also work through the problem and see how its own language may be confusing or actually incorrect. It also may create a sandbox of a project that can be workable but might not reflect the process taught in the classroom. Thus, we must complete the project before giving it to students and ensure that the language, instructions, and questions are all cohesive and oriented to the standards.

Without sustained professional development, AI integration risks following the path of earlier technologies: reduced to time-savers, detached from deeper pedagogy, or limited to repackaged content. If uncritically adopted, these tools may reinforce standardization rather than disrupt it.

The use of textbook curriculum is preferred over outsourcing curriculum creation to AI without critical review, as well-designed textbooks still offer intentional scope and internal consistency. To realize AI's potential as a tool for design-based instruction, educators need systems that support experimentation and reflection. This includes:

- Ongoing training in prompt design and task adaptation
- Time for collaborative planning and iterative testing
- Professional learning communities to analyze and refine AI-generated tasks
- Guidance on documenting and sharing classroom adaptations
- A shift in mindset that treats AI not as an end-product generator but as a co-author of inquiry-based, context-rich learning

Educators must be positioned as active designers, and reorient the pedagogical questions in a new way not afforded to us as practitioners: as an active lever in our classrooms.

9 Conclusion

AI in education is often framed around speed, automation, or the threat of academic dishonesty. This paper has argued for a different paradigm to treat generative AI as a partner in shaping relevance, creativity, and critical thinking in mathematics education. With this approach we can engage students in a new, dynamic mathematical understanding that goes beyond rote procedure or irrelevant word problems.

When guided by intentional pedagogy, tools like GPT-4 can help teachers create math problems that are resonant with the social and historical context we live in. There is a challenge for educators to review problems to ensure their internal cohesion, their relevance to the standards, and technical accuracy, but I believe this practice makes us sharper as educators, by reflecting on the needs of our students and also having a way to quickly mitigate any issues with further prompting of GPT-4 or other AI models.

Used with intention, AI can help educators reconnect math to the living world, not just to reflect it, but to interrogate it, reimagine it, and create avenues for students to understand and fight. Math is

not neutral, and AI allows for teachers to answer that universal question “why do I need this?” in the development of these types of problems.

This paper offers one example of what that process can look like. Many more are needed. As a full-time teacher in Detroit, this work was written in the margins of daily life, between teaching, parenting, and taking graduate classes. That is precisely what makes AI promising. It can support education by creating avenues for us to impact our classrooms with intention and flexibility.

References

Abah, J. A. (2017). Viewing basic math through the lens of history: Undergraduates' reflective learning in a history-augmented mathematics classroom. *Waikato Journal of Education*, 22(4), 33–48.

Aguirre, J. M., & Zavala, M. del R. (2013). Making culturally responsive mathematics teaching explicit: A lesson analysis tool. *Pedagogies: An International Journal*, 8(2), 163–190. <https://doi.org/10.1080/1554480X.2013.768518>

Buck, E. L., & Damico, M. M. (2021). Teacher learning about equitable participation in math class: Iterative analysis of video and narrative. *Teaching for Excellence and Equity in Mathematics*, 12(1), 15–25. <https://journals.charlotte.edu/teem/article/view/1437>

Hoyles, C., Noss, R., Kent, P., & Bakker, A. (2013). Mathematics in the workplace: Issues and challenges. In A. Damlamian, J. F. Rodrigues, & R. Strässer (Eds.), *Educational interfaces between mathematics and industry: Report on an ICMI-ICIAM study* (Vol. 16, pp. 43–50).

Hwang, W.-Y., & Utami, I. Q. (2024). Using GPT and authentic contextual recognition to generate math word problems with difficulty levels. *Education and Information Technologies*, 29, 1–29.

Lesh, R., Doerr, H. M., Carmona, G., & Hjalmarson, M. (2003). Beyond constructivism. *Mathematical Thinking and Learning*, 5(2–3), 211–233.

OpenAI. (2025). *ChatGPT (GPT-4.1)*. <https://chat.openai.com/>

Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? *Review of Educational Research*, 82(3), 330–348. <https://doi.org/10.3102/0034654312457429>

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM - Mathematics Education*, 52(1), 1–16. <https://doi.org/10.1007/s11858-020-01130-4>



Mosum Trivedi is a mathematics teacher at Detroit Public Schools and a Masters candidate at Wayne State University. She teaches Geometry and IB Math SL Applications and Interpretations, with research interests in mathematics education, culturally relevant pedagogy, and AI integration in teaching.