2025

AI FOR EDUCATION
&
STUDENT ACHIEVEMENT PARTNERS

A Guide to Integrating Generative AI for Deeper Math Learning

AI for Education

Guide to Integrating Generative AI for Deeper Math Learning

This document represents an initial exploration of how to harness the power of generative artificial intelligence (GenAI) to enhance, not replace, the cognitive lift and meaningful learning in K-12 math classrooms. Developed with an understanding that both the technology and its educational applications are rapidly evolving, this document explores practical strategies for using GenAI to support students' productive struggle while maintaining the essential cognitive work that drives learning. We offer these insights as a conversation starter and support as educators continue to wrestle with how GenAI can serve instructional needs and how content and pedagogy intersect with this evolving technology.

Goals of This Resource

While ChatGPT and other GenAI tools are at the earliest stages of their development, recent **research** shows teachers have begun exploring ways to integrate them into planning and classroom instruction. Leaders and teachers are grappling with determining how to integrate GenAI in a way that effectively supports learning and strengthens student outcomes at a time when there is little evidence or research to guide their approaches.

Leveraging our collective expertise in the current uses of GenAI in math classrooms and high-quality math instruction, AI for Education and Student Achievement Partners have created this guide focused on how mathematics educators can begin to use these tools to enhance, not replace, students' cognitive lift in the classroom.

This guide is a work in progress. We aim to gather input, learn alongside our teacher and instructional leader community, refine our understanding, and amplify best practices. We will be seeking feedback from classroom teachers who apply these strategies so that we can collectively better understand the impact on student learning.

Key Assumptions

- This resource assumes teachers and students have a basic foundation in AI literacy. The suggested activities serve a dual purpose: supporting both math instruction and the development of AI literacy skills.
- This resource also assumes a shared definition of high-quality math instruction that encompasses essential mathematics, connected mathematics, meaningful mathematics; and student-centered mathematics. For more on this topic, see the e2 Instructional Practice Framework™.

Definitions of Key Concepts

- **Mathematics instruction:** instruction that is focused on developing students' understanding of mathematical concepts, proficiency with procedural skills, and demonstration of problem-solving abilities.
- **Cognitive lift:** mental effort required to engage with content while learning to demonstrate deep understanding and critical thinking skills.
- **Productive struggle:** meaningful interactions with high-quality math content that poses challenges through which to learn.
- **Counterproductive struggle:** challenges with content or tasks that create barriers to learning.
- **AI literacy:** knowledge, skills, and mindsets that enable individuals to use AI safely, ethically, and effectively.
- **Generative AI:** subset of AI that can generate new content such as text, code, images, audio, and video. Generative AI models are developed using massive amounts of data, primarily from the internet, which allows them to recognize patterns and predict what content might come next in a sequence.
- **Large Language Model (LLM):** an artificial intelligence system trained to recognize patterns in large amounts of text data and replicate those patterns to generate human-like text. LLMs can perform language-based tasks such as writing essays, summarizing texts, and translating by calculating the most probable response to a text or voice input.
- **Chatbot:** AI tool designed specifically to simulate conversation with human users through text or voice interactions, often powered by LLMs.

Productive (and Counterproductive) Struggle in Math Classrooms

When math instruction is structured to focus on procedural skills and memorization, absent conceptual understanding or opportunities for student agency in selecting problem-solving methodologies, students are more likely to internalize the beliefs that there is one right approach to a solution, that mistakes are failures, and that peers are competitors. When students approach math with this mindset, they are more likely to engage in counterproductive struggle, which increases math anxiety and hinders positive development of math identities.

Alternatively, when math instruction is focused on building conceptual understanding, providing multiple pathways towards a solution, fostering student agency, and leveraging student assets in approximations and partial solutions, students are more likely to deepen engagement and build positive math identities.

> **The concept of "productive struggle" is often articulated in education circles, but it is not frequently enough attended to in terms of answering these questions:**
>
> - **What is a productive struggle that supports content-specific learning?**
> - **Conversely, what is counterproductive struggle that negates or detracts from learning?**

Building on the core components of high-quality mathematics instruction, the following chart illustrates how productive and counterproductive struggle manifest in math classrooms. For each key area, specific examples demonstrate the crucial difference between challenges that advance learning and those that may impede it. This chart serves two essential purposes:

1. It provides concrete examples of how productive and counterproductive struggle appear within each component of high-quality math instruction as a foundation for the GenAI use cases outlined below.
2. It serves as a practical tool for teachers and leaders to reflect on and evaluate struggles in their own math classrooms, helping them ensure their instructional choices promote meaningful learning.

High-Quality Mathematics Instruction	Productive Struggle	Counterproductive Struggle
Essential Mathematics: Engage in cognitively challenging mathematics that focuses on grade-level or course-level content standard(s)	Tasks offer multiple entry points and scaffolds when necessary to access the grade- or course-level content. Tasks allow for multiple pathways so that students can demonstrate understanding of grade-level mathematics concepts in a variety of ways.	Scaffolds that “make sense” of the problem for students (e.g., providing prompts that break down the steps of a multi-step word problem, providing fill-in-the-blank number sentences that represent the problem).
Connected Mathematics: Relate new content to math content within or across grades that builds towards future mathematical concepts; build procedural skill and fluency from conceptual understanding	Tasks leverage the coherence of math concepts across and within grades and courses. Instruction builds on students’ prior math knowledge, ideas, assets, and goals as they develop grade-level and course-level understanding. Students share their reasoning in multiple ways, including using multiple representations.	Practice is of procedural skills or fluency, absent conceptual understanding. Instruction prioritizes only answer-getting and/or solving a problem in one way.
Meaningful Mathematics: Engage in mathematical reasoning, discourse, communication, problem-solving, and modeling while engaging in the math content; position students as the mathematical experts, thinkers, and doers in the classroom	Tasks center on current events relevant to students’ lives and that can be analyzed mathematically. Tasks connect with the real world and provide opportunities for students to develop skills and knowledge relevant to their academic and professional goals.	Supports “simplify” the problem (e.g., removing context, highlighting key words/phrases) as a way of making the problem more accessible. Tasks do not include data or contexts relevant to students and their communities when they are exploring mathematical concepts.

	Supports “simplify” the problem (e.g., removing context, highlighting key words/phrases) as a way of making the problem more accessible.	Students are forced to only make meaning within a problem’s provided context (such that students’ identities, communities and cultures are not valued).
Student-Centered Mathematics: Engage in mathematical reasoning, discourse, communication, problem-solving, and modeling while engaging in the math content; position students as the mathematical experts, thinkers, and doers in the classroom	Instruction provides access to math tools and offers ample time to explore math concepts. Students co-construct, share, and own the mathematical ideas produced in the lesson. Students build on one another’s shared ideas and math knowledge and ask questions about each other’s thinking. Instruction positions multiple languages and cultures as valuable assets.	Scaffolds/supports select the tools for students to use, provide the representation for students to use, and/or dictate the steps that students will take. Teachers are the only voices summarizing the lesson takeaways. Instruction provides little or no opportunities for student discourse; students are discouraged or prohibited from utilizing home and everyday language to learn mathematics.

Enhancing Math Instruction with GenAI

GenAI can be viewed as a tool for extending thinking, similar to the view of calculators in mathematics instruction. Over time, norms and standards for using calculators in ways that support meaningful engagement with computation have been developed. Just as math instruction must be carefully designed to position the calculator as a tool to support productive struggle and not as a shortcut, GenAI tools must be integrated with similar intentionality—considering how these tools can enhance rather than replace student access to productive struggle.

Specialized AI Tools for Mathematics

While GenAI tools show definite promise in enhancing mathematics education, our field tests have revealed significant limitations[1] in calculation reliability[2] and accuracy, particularly in general-purpose (foundational) models like ChatGPT. These Large Language Models (LLMs) operate by predicting the next word in response to a prompt, not by performing actual calculations.

> **Therefore, it is our strong recommendation that educators use specialized tools such as Khanmigo or Desmos for any task involving computation.**

These tools have been developed with additional fine-tuning to support numerical calculations beyond the capabilities of foundational LLMs. They also typically have structured interfaces that support the input of mathematical symbols and equations. Foundational LLMs are most useful for language-based tasks such as lesson planning, conceptual exploration, and generating alternative explanations and examples for students.

[1] https://www.forbes.com/sites/johnwerner/2024/10/07/ai-is-usually-bad-at-math-heres-what-will-happen-if-it-gets-better/

[2] https:/arxiv.org/pdf/2503.21934v1

AI Tool Categories and Their Optimal Uses

Foundational Large Language Models

Foundational LLMs such as ChatGPT, Claude, Google Gemini, and Microsoft Copilot are most suitable for language-based tasks, and should never be relied upon for computational accuracy. These models excel at:

- **Brainstorming:** generate alternative teaching strategies, creative ways to address standards, or metaphors to clarify abstract mathematical concepts. They can also act as a conversational thought partner, helping educators explore different instructional approaches, refine ideas through back-and-forth dialogue, and consider perspectives they might not have otherwise thought of.
- **Designing innovative projects and assignments:** generate ideas or outlines for projects, tiered assignments, or performance tasks that integrate mathematical thinking with real-world applications. They can suggest themes, prompts, or scaffolding structures based on standards and learning goals.
- **Drafting student-friendly communications:** draft clear definitions, age-appropriate learning objectives, and simplified explanations of complex terms or procedures. This can help make math content more accessible and relatable for students.
- **Identifying common misconceptions:** list typical errors or partial understandings that students might have with a given concept (e.g., confusing the numerator and denominator in fractions), along with suggestions for addressing these issues through targeted instruction or scaffolding.
- **Translating:** translate word problems or teacher communications such as emails, newsletters, and assignment instructions into other languages. Note- it is recommended that a native speaker review the output whenever possible to ensure cultural and linguistic accuracy

Specialized Tools

Specialized AI-powered math tools are purpose-built applications that integrate AI capabilities specifically for mathematical tasks. While powered by LLMs, these tools receive additional fine-tuning during the development process that allow them to perform actual calculations rather than merely predict text patterns. Khanmigo, for example, incorporates a calculator in its operations rather than relying solely on LLM outputs. These tools have received additional training specifically on mathematical concepts, and often contain specialized interfaces that can accommodate symbols, notation, and structured equations. **For any task involving calculations, specialized math tools should be the preferred choice for both teachers and students.**

Disclaimer: *the tools listed below are provided only as examples of AI-powered math tools. This list is not exhaustive, and inclusion of a tool does not constitute endorsement. We encourage educators to explore a range of tools to find those that best fit their instructional goals and students' needs.*

- **Khanmigo:** an AI-powered tutor developed by Khan Academy, integrated with ChatGPT-4. Offers conversational, Socratic-style guidance to help students think through problems rather than simply providing answers.
- **Mathway:** provides step-by-step solutions to a wide variety of math problems, from basic arithmetic to calculus. Supports photo uploads of math problems.
- **Photomath:** mobile app that supports photo uploads of handwritten or printed math problems. Provides animated, step-by-step explanations.
- **Gauthmath:** AI-powered step-by-step problem solver with human tutor support when needed. Supports photo and PDF uploads of math problems.
- **Snorkl:** formative assessment tool that analyzes student work, identifies misconceptions, and provides actionable insights for teachers. Allows students to insert voice recordings of their thinking as they work through problems on the platform.
- **Mathos AI:** provides both solutions (Math Solver) and step-by-step AI tutoring (Math Tutor). Supports photo and PDF uploads, as well as voice inputs.
- **Wolfram Alpha:** an advanced computational engine that solves equations, visualizes data, and integrates math with science, engineering, and other fields to explore real-world math applications. Provides a wide variety of practice problems.
- **Desmos:** AI-powered online platform offering a suite of tools for teacher planning, diagnostic assessment, interactive practice, data visualization, and graphing.
- **Julius AI:** a generative AI assistant for analyzing and interacting with data. Allows teachers to ask questions of data sets directly using natural language.

Key Principles for GenAI Integration

When leveraging AI tools in math instruction, consider the following:

- **Tools should support, not circumvent, productive struggle for students.**
- **Integration should enhance, not replace, proven instructional practices.**
- **Usage should align with students' developmental readiness and mathematical learning goals.**
- **GenAI should augment educators' pedagogical expertise, content knowledge, and knowledge of students.**
- **LLMs should be used for language-based tasks like conceptual exploration, planning, and generating examples.**
- **Specialized, math-specific tools should be used for tasks requiring computation.**

A way to consider meaningful use of GenAI that supports student learning is to move away from activity- and technology-driven tasks, and instead focus on student learning. Ask:

1. **What are students learning through this task?**
2. **Where is productive struggle essential to student learning?**
3. **Where might GenAI benefit learners or learning outcomes? Where could it mitigate counterproductive struggle?**

A Note About GenAI Use for Elementary Teachers

- At the elementary level, GenAI tools should be used primarily by teachers for lesson planning, preparation, and selective modeling. This use aligns with elementary students' developmental stages—when children are still building foundational cognitive skills like causal reasoning, source evaluation, and metacognitive awareness. These developmental factors make it premature for young students to independently evaluate AI-generated content or critically assess its limitations.

- Until there is compelling evidence of positive impacts on learning through direct student interaction with GenAI tools at the elementary level, as well as adequate safeguards in place to completely eliminate the risk of exposure to inappropriate content, elementary teachers are advised to restrict GenAI use to their own workflows. This is an evolving field, and while there may be some direct interaction that is supportive of students, we encourage caution here.

- Many GenAI tools impose age restrictions that prevent elementary students from engaging with them directly. Those that do allow elementary use require close oversight and adult guidance to help students interpret responses critically.

Strategies for Using GenAI to Support Deeper Math Learning

- Following are examples of effective and ineffective applications of GenAI, including worked chatbot transcripts that illustrate these practices.

- Reminder: Use specialized tools as described above for any tasks involving calculations, equation solving, or mathematical modeling. General-purpose language models like ChatGPT or Claude are not designed to perform reliable mathematical computations and may produce incorrect results.

- Within the following tables, student use cases are highlighted. Examples that assume direct student use **are targeted to middle and high school students with parental permission.**

For Students

Use Case	Detailed Description	High-Quality Instructional Support	Cautions & Considerations
Engage with AI Tutoring Worked example with computational errors Use specialized tools	AI math tutoring tools have been specially developed to help the students think through problems, asking questions and providing feedback to guide thinking without providing answers directly.	• Positions AI as a tool to help students think through problems and reveal various entry points to reinforce underlying mathematical processes • Provides private forum for asking questions without fear of judgment • Offers 24/7 access to feedback and support	• Avoid using general-purpose language models such as ChatGPT. Use specialized tools (see list above). • Recognize that responses are limited to the tool's interface and training data. • Structure use of AI tutors to encourage self-reflection and gradual skills release; strive for students to shift to self-checking after mastering the concept under study to prevent over-dependence. • Ensure the AI tutor is aligned with specific curriculum to avoid conflicts
Conduct error analysis and feedback Use specialized tools	AI can check completed student work for accuracy, provide immediate feedback, and answer questions, guiding students to a clear understanding of where and how errors occurred, as well as how to correct them.	• Scaffolds conceptual understanding by targeting specific areas of breakdown • Provides a safe space for students to identify gaps in their understanding and ask clarifying questions without fear of judgment	• Avoid using general-purpose language models such as ChatGPT. Use specialized tools (see above). • Recognize that responses are limited to the tool's interface and training data.

For Students (continued)

	AI can generate a step-by-step explanation of a solution to a math problem and prompt students to understand the process behind an answer. AI can prompt students to do an error analysis of a worked example (with errors).	• Provides opportunities for students to explore multiple explanations of mathematical concepts that can be easily tailored to their interests and experiences • Allows for analyzing drawings and handwritten equations through multimodal functionality.	• Note risk of cognitive offloading and over-dependence; structure AI interactions in the manner of tutoring sessions by explicitly prompting the model not to provide direct answers, but rather to guide students' thinking to determine the correct solution themselves when they get stuck
Practice with procedural skills Use specialized tools	Students can scale up their practice with procedural skills by using AI to generate personalized problems related to the current skills under study.	• Increases access to practice opportunities for specific skills • Invites scaffolded practice that build on prior math knowledge and/or skills and connect to grade-level content • Allows students to practice in a time and place of their choosing • Allows students to co-construct meaningful practice problems by customizing the content to their background knowledge and interests	• Avoid using general-purpose language models such as ChatGPT. Use specialized tools (see above). • Recognize that responses are limited to the tool's interface and training data. • Note risk of cognitive offloading and over-dependence; structure AI interactions in the manner of tutoring sessions by explicitly prompting the model not to provide direct answers, but rather to generate relevant practice problems and guide students' thinking to determine the correct solution themselves when they get stuck.

For Teachers

Use Case	Detailed Description	High-Quality Instructional Support	Cautions & Considerations
Create lessons, materials, and interventions Worked example	AI can serve as a thought partner to help teachers develop lesson plans, instructional content, and targeted interventions or extensions by suggesting activities aligned to learning objectives and filtered through the teacher's professional judgment.	• Adds capacity to the lesson planning process • Ensures alignment of interventions and extensions to primary learning objectives • Generates suggestions for diverse examples and scenarios connected to real-world math applications • Redesigns instruction and assessment to incorporate math practices such as mathematical reasoning and communication	• Carefully screen AI-generated content for bias and developmental appropriateness • Since lack of contextual awareness can lead to misalignment with adopted curricula and department/school/district goals, ensure these are in alignment before deploying AI-generated content with students
Generate authentic practice problems Use specialized tools	AI can create highly contextual practice problems at volume, incorporating real-world scenarios, student interests, and standardized test referents.	• Quickly creates engaging, relevant practice problems at scale • Generates complex problem sets that provide multiple pathways to demonstrate understanding • Creates coherent assessments aligned to reference documents like curricula, standards, and standardized test exemplars	• Always check AI-generated problems for accuracy and solution viability • Guard against bias and cultural misrepresentation in AI-generated word problems • Ensure that AI-generated problems are in alignment with formal curriculum and learning objectives

For Teachers (continued)

Provide differentiation/ Individualized instruction Worked example	AI can assist with developing modified assignments, alternative assessment methods, and differentiated teaching strategies/explanations based on specific student learning needs.	• Creates materials with multiple entry points for diverse learners • Facilitates efficient scaffolding of instruction for a variety of readiness levels and backgrounds • Brainstorms varied approaches to content presentation and alternative teaching strategies, examples, and explanations to facilitate student understanding	• Ensure that modifications maintain grade-level-appropriate task complexity and rigorous learning expectations • Always check AI-generated problems for accuracy and solution viability • Ensure that AI-generated materials are in alignment with formal curriculum and learning objective
Break down tasks and concepts Worked example	AI can help unpack challenging tasks and concepts by generating related content such as prerequisite skills, warm-up questions, common errors or partial understandings, helpful interventions, and other pedagogical strategies.	• Assists with activating students' prior math knowledge by planning activities that activate prerequisite skills • Anticipates and proactively plans to address common errors and partial understandings • Brainstorms lists of hints and alternative explanations for specific problem types that do guide students who are stuck without giving away answers directly • Offers insights of how to break down complex problems into manageable chunks to facilitate understanding	• Ensure that AI-generated content is accurate, aligned to learning objectives and established trajectories, and free from bias

For Teachers (continued)

Analyze student data and student work Use specialized tools	AI can analyze student data and/or student work to identify trends and action items related to topics such as student thinking, common errors and partial understandings, areas of strength, the effectiveness of a given teaching strategy, student interests, and more.	• Assesses student performance on close-ended assessment tasks and identifies trends/action items • Efficiently scores constructed response data, including drawings • Identifies and awards partial credit for correct thinking in some areas of an assessment task, providing deeper insight into student understanding	• Risk of compromising student data privacy; never enter students' personally identifiable information into an AI tool; always anonymize data before uploading • Be aware of biases inherent to the training data that can be reflected in action items (e.g., small sample sizes for marginalized student populations can lead to overestimation or underestimation of their performance)
Create graphs, charts, and interactives Use specialized tools	AI can generate visual representations of data and become an interactive digital tools to illustrate mathematical concepts for instruction and assessment	• Enhances concept visualization • Supports multiple representations of mathematical concepts and data • Supports a wider variety of instructional content creation	• Ensure that digital interactives are accessible to all students, including those using assistive technology such as screen readers or alternative input devices • Verify accuracy and alignment to curriculum and learning objectives

Who We Are

AI for Education is committed to leading the responsible adoption of Generative AI in education by providing comprehensive AI literacy training and AI guidance and policy development to schools and educators worldwide. As a pioneering organization that has partnered with some of the largest school districts across the U.S., they develop and publish a wealth of free, practical, and innovative resources while working with academic institutions to conduct training, develop curriculum, define policy, and implement customized AI solutions. With the goal of providing AI literacy training to 1 million educators, their mission focuses on empowering teachers to expertly navigate AI technology, ultimately improving student outcomes and preparing them for the future.

aiforeducation.io

Student Achievement Partners (historically known as Achieve the Core), is a central voice in the conversation about high-quality instructional practices and materials in K12 Mathematics and Literacy in the US. Founded in 2011, members of the original SAP team were the lead writers of the Common Core State Standards and began developing tools and resources designed to give more students access to high-quality, standards-aligned math and literacy instruction. As a team of former educators, SAP understands that the most useful and effective tools and guidance are grounded in both the latest pedagogical research and educator insights into what works in the classroom. In the last four years, SAP has developed an expanded framework for instructional excellence, e .

learnwithsap.org