Productive Struggle

How Artificial Intelligence Is Changing Learning, Effort, and Youth Development in Education

By Amy Chen Kulesa, Marisa Mission, Michelle Croft, and Mary K. Wells

JUNE 2025

Bellwether

CONTENTS

Introduction

It is a Tuesday morning in early November. In a ninth-grade Language Arts class, Ms. Lopez moves between desks as students craft a science-fiction story set in the year 2050. She kneels beside Mateo, who sits in front of an artificial intelligence (AI) writing tool. "I have ideas," Mateo whispers, "but the words won't come out." A few feet away, Jada toggles between her notebook and the AI tool that generates quirky what-ifs. Each suggestion sparks a fresh question, a scribble, and a playful mash-up that Jada weaves together. Across the room, Laila copies and pastes the first written idea she generated with the same AI tool. The writing is polished and error-free, yet when Ms. Lopez asks a follow-up question, Laila struggles to explain the reasoning in her own words. In just a few minutes, the same AI tool has enabled divergent outcomes: one student immobilized by the blank page, one invigorated with curiosity, and another bypassing the challenge altogether. These tensions, between motivation and disengagement, support and shortcut, creativity and compliance, are not new. But AI reshapes and magnifies them in subtle ways that demand urgent attention.

Scenes like this unfold in classrooms every day, with or without AI. Yet, today's stakes are high. According to the 2024 National Assessment of Educational Progress, 40% of fourth graders and 33% of eighth graders performed below basic in reading, while 24% and 39%, respectively, did so in math.[1] COVID-19 pandemic disruptions compounded long-standing inequities,[2] and as federal relief funds wind down,[3] enrollment declines,[4] and the federal landscape shifts,[5] students and teachers face mounting pressure with fewer resources.

Into this moment enters generative AI (GenAI),[6] a tool that promises efficiency and customization for teachers and students alike but also carries risks of dependency and detachment. When deployed well, AI could reduce administrative burdens, personalize supports, and generate insights that guide strong relationships.[7] But when used indiscriminately, it may erode cognitive effort, weaken instructional judgement, and displace the very relationships that fuel learning.

AI is rapidly advancing (Sidebar 1). Yet, its development does not override cognitive science and pedagogical research showing that students learn when they are challenged, supported, and given opportunities to reflect. This dynamic, often called "**productive struggle**," remains fundamental in learning. When students engage in tasks that are just beyond their current mastery, supported by timely feedback and opportunities to iterate, they build knowledge, resilience, and agency. At the same time, AI invites a revisit of what productive struggle should look like in a technology-rich world. Not all friction may be inherently beneficial, and not all ease may be harmful. In some cases, AI may reduce surface-level barriers, such as organizing initial ideas or decoding, freeing up students to spend more time exploring, revising, and persisting. AI-facilitated ease may unlock curiosity, extend time-on-task, or enable students to reach cognitive depths they may not previously have accessed.

Rather than asserting that learning must be a certain way, the better question becomes: **when does ease enable greater learning, and when is ease a shortcut with a hidden cost?** The answer may vary by age, developmental stage, content area, prior knowledge, motivation, and relationships. It may also depend on the context in which learning occurs, whether students feel safe, supported, and capable of taking risks.

Teachers seek to calibrate this balance by scaffolding questions, pacing instruction, and offering "just-right" guidance to help students navigate complexity without becoming overwhelming. AI changes the dynamics of that orchestration, raising new questions for educators, system leaders, and tool developers: how much cognitive effort should AI alleviate, and how much must it intentionally preserve? What guardrails ensure that adaptive supports do not drift into over-scaffolding? And how does this technology evolve what comprises the critical skills necessary for human development?

Addressing these questions cannot fall solely on the shoulders of students or teachers. It is human nature to favor ease, and many tools — especially those designed and incentivized to scale quickly — are built to please users, not necessarily to preserve what is most essential for long-term and holistic student development. While schools and system leaders play a critical role, the complexity of the market and the pace of technological change make it difficult to know where responsibility ultimately lies, raising the risk that this becomes a hidden problem without a clear owner.

This report aims to change that. It moves beyond polarized debates of "is AI good or bad?" and instead dwells in the murkier, more consequential space where nuance lives. By weaving together evidence from the science of learning, capabilities of emerging technology, and early empirical research, this report explores the blurry boundaries where AI can amplify effective teaching and learning, and where it risks undercutting them.

The goal is not to pick sides; rather, it is to illuminate the design, research, and implementation choices that will determine how and whether AI eases or impairs the kind of productive struggle that cultivates lifelong learners. Students' futures hinge not just on their ability to prompt, produce, and retrieve, but on their ability to think critically, engage, and discern. Collectively, education leaders, funders, policymakers, and researchers must hold the tensions and center students like Mateo, Jada, and Laila, whose futures will be shaped not just by the tools they use, but also by how and why they use them.

SIDEBAR 1

Recent Developments in AI

Recent advancements in AI mark a significant shift in the technology's capabilities, with major implications for education. Advanced reasoning models employing chain-of-thought reasoning have emerged as a new standard,[8] allowing AI to tackle more complex and higher-order tasks and generate nuanced and accurate outputs, all without needing increased computational resources.

Concurrently, agentic systems — composed of autonomous AI agents capable of acting independently — have rapidly evolved, enabling applications that autonomously execute tasks, interact with computer systems, and even collaborate to solve problems.[9] The integration of advanced reasoning and agentic capabilities has also birthed Deep Research models[10] capable of performing complex, knowledge-based tasks with near-expert levels of accuracy.[11] These technologies, particularly when combined with robotics and multimodal capabilities (e.g., vision[12] and voice), will transform educational practices, though the concrete benefits or drawbacks are not yet clear.

Additionally, China has emerged as a major player in AI with cost-effective training innovations that enabled DeepSeek's R-1 model.[13] The resulting pressure and global competition will create strategic implications for education in terms of national security and workforce preparedness. As AI rapidly progresses toward even more sophisticated capabilities — including the eventual advent of artificial general intelligence[14] capable of performing intellectual tasks on par with humans — both the challenges and opportunities of mindfully and responsibly integrating AI into education will continually evolve.

From Struggle to Mastery

WHAT THE SCIENCE SAYS

Although known by different names throughout the literature (e.g., desirable difficulties,[15] zone of proximal development[16]), **productive struggle** generally refers to "the process of engaging with challenging tasks or problems that require effort, critical thinking, and persistence to solve," and typically includes running into obstacles, making mistakes, and experiencing discomfort — all while still working toward a solution.[17] Notably, no matter the name, researchers have identified that there must be an element of appropriateness.[18] In other words, the struggle must be *productive*, the difficulty must be *desirable*, and the zone of development must be *proximal*; the task should be something the student may not be able to do independently but can reasonably accomplish with support. Struggle for struggle's sake can deter learning.[19] However, when appropriately tailored to the student's capability level, the struggle can enhance the cognitive processes critical to learning.

While there are many cognitive factors that contribute to learning, this first section of the report focuses on four broad components: **memory and information processing**; **attention and engagement**; **motivation and mindset**; and **metacognition and self-regulation**. These components have been identified by psychologists who, over the past century, have investigated what learning is and how to amplify it. They were driven largely by major political events and technology developments (e.g., the Cold War, the rise of personal computing), as well as the resulting need for people who can understand and solve complex problems.[20] The current moment is remarkably similar: The U.S. is once again faced with major political change, technological milestones, and the need for students who can solve even more complex problems. A review of what cognitive science has already identified as the critical components of learning provides a foundation for understanding how AI might amplify or weaken students' future skills.

A review of what cognitive science has already identified as the critical components of learning provides a foundation for understanding how AI might amplify or weaken students' future skills.

Memory and Information Processing

What Is It?

One reason why AI, despite being just a very advanced computer, can mimic human performance is that at its core, human cognition is an information-processing system — albeit an incredibly intricate, complex, and fast one. When humans encounter new information, it flows through three states of memory: from sensory, to working, to long-term memory.[21] The process of learning sits between the working and long-term memory stages, during which information is organized, connected to other pieces of information,[22] and encoded into schemas of prior knowledge that students will reference in the future as they absorb new pieces of information.[23]

How Does Productive Struggle Enhance It?

The process of encoding information into long-term memory is difficult and effortful in part because the working memory has limited capacity and duration.[24] If there is too much new information to encode, or the new information is too complex, then some of the information is lost (i.e., not learned). However, there still needs to be some engagement, which often means struggle: research shows that actively processing information through recalling it, organizing it, connecting it to existing knowledge, talking about it, and practicing it leads to stronger encoding and better long-term retention.[25] In this case, productive struggle occurs when the working memory has an appropriate amount of information to process and students are actively wrestling with that information to encode it for their long-term memory. Once the student has encoded the new knowledge, then the struggle becomes retrieving that information from long-term memory to use and apply it.[26]

Attention and Engagement

What Is It?

Attention is difficult to define[27] but can be thought of as the cognitive mechanism that selects which information to focus on and which information to ignore.[28] It is a prerequisite to learning: A student must pay attention to information in order to process and learn it.[29] Moreover, human attention is limited, and learning requires *sustained* attention (i.e., focus).[30] If a student's attention is divided (i.e., unfocused), the information may never fully enter working memory, and thus may not be learned.[31] Relatedly, engagement refers to ways of capturing and keeping students' focus.[32] The more engaged a student is with the classroom, educator, or material, the more attention and focus they are giving it, and the more information they will be able to absorb, process, and encode.

How Does Productive Struggle Enhance It?

The appropriate level of challenge is key to keeping students' attention; without enough of a challenge, learners become bored and lose focus. Psychologist Mihaly Csikszentmihályi proposed the concept of a "flow" zone, where a task's difficulty level is high enough to challenge a student's competencies, but not so high that they become discouraged and therefore disengaged.[33] Research has also shown that children can practice sustained focus and build their "cognitive endurance" by doing more challenging tasks in shorter bursts.[34] By calibrating a challenge so it is high but manageable, productive struggle places students in the flow zone — maximizing sustained attention without tipping into boredom or overload. Thus, struggle-driven tasks are powerful because they naturally capture and hold a student's focus.

Motivation and Mindset

What Is It?

Motivation can be defined as an internal condition that "arouse[s], direct[s], and maintain[s]" students' behavior toward learning goals;[35] motivation affects whether students *choose* to learn. While there are several related theories, two influential frameworks are the self-determination theory and the expectancy-value theory. In the former, motivation is either intrinsic or extrinsic.[36] Intrinsic motivators come from within a student, are linked to an inherent satisfaction with doing a task, and are associated with deeper engagement and learning.[37] Extrinsic motivators are not linked to inherent satisfaction and stem from expectations of the external environment; as a result, they are mainly useful in the short term.[38] In the expectancy-value theory, self-efficacy (whether an individual believes they have the skills to succeed at the task) and task value (how worthwhile a task appears to be) are the two factors that motivate a student to learn.[39]

How Does Productive Struggle Enhance It?

In the self-determination theory, a student's success after an appropriately challenging task produces a sense of accomplishment that boosts the self-satisfaction underlying intrinsic motivation.[40] Essentially, productive struggle increases an individual's intrinsic motivation to learn, creating a virtuous cycle of learning. In the expectancy-value theory, motivation plays more of a role in determining *whether* a struggle is productive: not only does increased self-efficacy and task value (and therefore higher overall motivation) lead to a greater likelihood of persisting through struggle, but self-efficacy is also related to a student's mindset.[41] Students with a "growth mindset," who believe they can grow their self-efficacy, are more likely to see a struggle as productive.[42] In turn, they are not only more likely to persist, but also more likely seek out productive struggles and reap the associated learning benefits.[43]

Metacognition and Self-Regulation

What Is It?

Metacognition is the ability to think about and manage one's own learning. It involves planning how to approach a task, keeping track of progress, and reflecting on what did and did not work, while also being aware of how emotions, motivation, and the learning environment affect that process.[44] The skill involves two components: knowing about cognition (e.g., knowing "how to study" and where one's strengths lie) and self-regulation (e.g., being able to question oneself through reflection).[45] Metacognition is key to students' ability to accurately calibrate their knowledge and take a strategy they learned in one context and apply it in a new context.[46] Students with strong metacognitive awareness tend to be better at judging what they do and do not know, which means they can both leverage their prior knowledge more effectively and study more efficiently by spending time on the material they have not mastered.[47]

How Does Productive Struggle Enhance It?

When students face a challenging problem, it forces them to slow down and become more aware of their own understanding (or lack thereof). Research indicates that struggle helps students learn more effectively by prompting self-questioning and strategy adjustment.[48] Concurrently, students with higher metacognitive skills may also recognize when their struggle is productive, understand the benefits of persevering, and choose appropriate strategies to navigate a challenge.[49] Students might think, "This is hard, but I can figure it out or learn from it," rather than, "This is impossible, I give up."

These four cognitive components of learning do not exist in isolation; each process or skill contributes to and reinforces the others. For example, engaging students in tasks that involve productive struggle requires sustained attention and engagement, which in turn facilitates effective memory and information processing by ensuring the student actively wrestles with new information, enhancing its encoding into long-term memory. Successfully overcoming appropriately challenging tasks boosts students' motivation and mindset, fostering greater self-efficacy and intrinsic motivation for future learning challenges. Throughout this process, students practice and develop metacognitive and self-regulatory skills as they become more aware of their own understanding, adjust strategies, and reflect on their learning experiences. In other words, enhancing one cognitive component through productive struggle naturally supports growth across the other components.

Given the role of productive struggle in boosting cognitive skills, AI's role in accelerating or weakening learning largely rests on how well it can turn the dial of productive struggle up or down to maximize students' cognitive activity.

The Possibilities

AI'S ROLE IN SCALING PRODUCTIVE STRUGGLE

Productive struggle enhances learning by amplifying cognitive processes, but what counts as "productive" effort varies by individual students. High-quality educators can often recognize when students are overwhelmed, struggling productively, or need more challenge, and can differentiate instruction or assignments accordingly. However, scaling differentiation is difficult when educators lack capacity and materials — which is where AI can help. Given the role of productive struggle in boosting cognitive skills, AI's role in accelerating or weakening learning largely rests on how well it can turn the dial of productive struggle up or down to maximize students' cognitive activity. Below are some situations where AI models and AI-powered tools can either enhance or inadvertently hinder the cognitive processes essential for learning.

Cognitive Component	Scaling Productive Struggle	Scaling Problematic Practices
Memory and Information Processing	AI-powered tools could incorporate effortful memory strategies like retrieval practice (also known as forced recall),[50] spaced repetition (studying topics over time),[51] and interleaving (weaving multiple topics within a single learning session). For example, an AI "study buddy" could track a student's memory retention rates and dynamically schedule review sessions when they begin to forget critical content, thereby continually revisiting the encoding process over time to maximize retention.	Without appropriate safeguards, AI-powered tools can reduce the processing needed to encode and learn information by giving students the answer or a summary of information right away. In the short term, this will lead to superficial memorization rather than deep understanding. Over time, this dependency could diminish a student's capacity for independently organizing and integrating complex information into coherent mental models.
Attention and Engagement	AI-powered tools could dynamically adjust content difficulty and format to maintain flow-level engagement. For example, an AI-powered virtual learning platform might detect waning student attention and automatically introduce interactive simulations or problem-solving tasks to re-engage with them.	Poorly designed AI tools might prioritize immediate gratification and entertainment, leading to superficial engagement. An AI-driven educational game that relies heavily on flashy visuals and quick rewards could distract students from deeper cognitive tasks, weakening their ability to sustain attention on less immediately stimulating, but crucial, learning activities.

Cognitive Component	Scaling Productive Struggle	Scaling Problematic Practices
Motivation and Mindset	AI-powered tools could personalize the level of challenge to nurture students' intrinsic motivation and cultivate a growth mindset through timely and intentional feedback. For example, an AI-powered tutor could celebrate a student's perseverance after tackling challenging problems or encourage continued engagement and intrinsic motivation by highlighting incremental progress rather than just correct answers.	If AI provides overly frequent reassurance and simplistic solutions, students might increasingly doubt their own abilities to independently overcome obstacles, fostering a fixed mindset rather than resilience and adaptability. AI could inadvertently harm a student's intrinsic motivation, sense of self-efficacy, and task value if they feel that the AI model did the real "work." Alternatively, students may begin to expect immediate success through AI assistance, gradually developing an extrinsic orientation toward learning.
Metacognition and Self-Regulation	AI-powered tools can be trained to integrate ways of building students' metacognitive skills by incorporating guided self-reflection or reality checks to calibrate understanding of their own skills. For example, an AI-driven writing tutor might ask students periodically to assess their comprehension and strategy effectiveness, building their capacity for strategic self-reflection and adjustment over time.	Extensive reliance on AI-provided guidance could discourage independent metacognitive engagement. Hypothetically, if AI consistently tells students exactly what to study and how to study it, and evaluates their comprehension automatically, students may gradually lose practice and confidence in self-directed learning — impairing their capacity to independently manage and reflect upon their learning processes.

A Peek Into the Future: The Imperative for Now

Often, when describing overreliance on AI, the topic at hand is cheating. While cheating has long been a concern in education,[52] the rise of AI tools has made it easier for students to engage in dishonest practices, such as passing off AI-generated essays or summaries as their own. Early state and district AI policies were grounded in preventing students from using AI to cheat,[53] but the policies have not dissuaded students from using AI in this way. In fact, it may be that students who were already likely to cheat are simply cheating in a new way: research indicates that student cheating behaviors remained stable in 2023, the year after the release of ChatGPT,[54] and AI detection software suggests that only 3% of assignments were generated by AI in 2024.[55] Given continuous developments in the technology[56] and greater awareness of AI,[57] there is continued concern that students may rely more on AI to cheat — but the emphasis is shifting. The risk is more than cheating; it is about students outsourcing the hard, mental work, like generating ideas or grappling with ambiguity, that builds their capacity to think independently.

Right now, there is a mismatch: Students are experimenting with AI tools,[58] while most school systems remain slow to adapt.[59] As a result, AI becomes a default collaborator, often shaping habits in silence. Some educators are thinking more about how to adjust assignments to reduce the chance that students will use AI.[60] Others are also shifting their concern from cheating to how AI may impact other skills needed for the workforce.[61] **Yet there has not been a cohesive call to investigate how AI can shift the cognitive processes that underpin learning**. There is more at stake than just dishonesty; students who become overreliant on AI tools that lack appropriate learning design are risking their cognitive abilities. For example, when students depend on AI to write their essays, they may miss out on developing essential skills such as brainstorming, research, critical thinking, analysis, and effective communication.[62] These are not one-and-done skills; they develop iteratively, through repeated attempts, across multiple contexts and subjects, with feedback and reflection over time. Without these formative experiences, students may miss the deep cognitive work that builds their capacity to think independently and adaptively — abilities students will need in the long term to navigate not just the workforce, but also life in general.

The emerging research on AI in education is far from conclusive. Studies focused on AI in K-12 are limited and leave many questions unanswered. While this report is not a comprehensive literature review, it aims to highlight illustrative examples from existing research that can be useful in understanding AI's potential role in productive struggle.

"The organization and critical thinking skills that must be employed when students write a longer, more formal piece are skills that will help students become better, more engaged citizens. The processes of brainstorming, researching, evaluating, selecting, analyzing, synthesizing, revising are all skills that help students become more critical citizens, more discerning consumers, and better problem-solvers."[63]

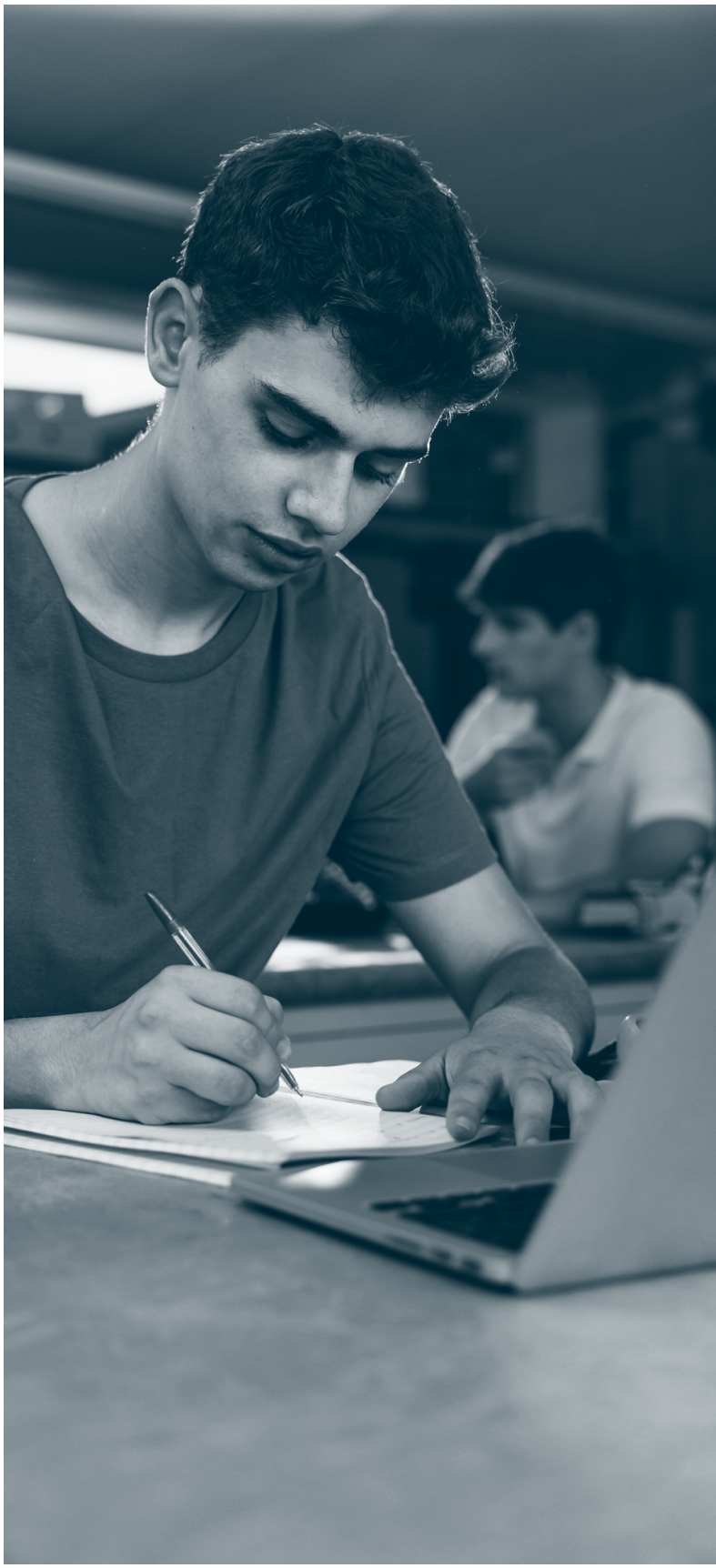
—Advanced Placement and National Writing Project Teacher, on the value of longer writing assignments in the digital world

Although a recent meta-analysis suggests that AI (specifically ChatGPT) can have a positive impact on learning performance,[64] **some studies suggest that AI tools may reinforce shortcuts instead of supporting deep learning**.[65]

For instance, in one study, undergraduates used either a traditional web search engine or AI to research a particular topic.[66] Students using AI experienced lower cognitive load, but at a cost. Although processing the information was easier, the quality of the students' final arguments was lower compared to students who used the traditional web search.[67] The researchers suspect that students may "not have engaged the deep learning processes as effectively as the more challenging traditional search task."[68]

Similarly, research has found that AI can increase short-term performance but not result in long-term learning. A study of ChatGPT access for high school math students found an increased short-term performance but worse long-term performance.[69] The researchers described "that students attempt to use GPT-4 as a 'crutch' during practice problem sessions, and when successful, perform[ed] worse on their own."[70] Likewise, when university students for whom English was a second language received ChatGPT support in a writing task, the ChatGPT group had greater improvement in the essay score, but there were no significant differences in knowledge gain and transfer.[71] This finding led researchers to caution the potential for "metacognitive laziness," which they defined as "learners' dependence on AI assistance, offloading metacognitive load, and less effectively associating responsible metacognitive processes with learning tasks."[72]

These challenges are exacerbated as students use AI to offload thinking instead of supporting thinking and learning. A study of how university students use Claude found that almost half (47%) of the student-AI conversations were "direct," which means the student was looking for answers with limited engagement.[73] These interactions are causing some experts to hypothesize that users will favor AI over engaging in meaningful learning.[74]



It does not have to be this way. AI tools, when intentionally and thoughtfully designed, can enhance learning for both students and teachers rather than hinder it.

It does not have to be this way. AI tools, when intentionally and thoughtfully designed, can enhance learning for both students and teachers rather than hinder it (Sidebar 2). For example, in the same study of high school students with access to ChatGPT, a subset received access to a different version of ChatGPT-4 that was directed to act as a math tutor and refuse to give specific answers.[75] The tutor version of ChatGPT instead prompted students to recall information and problem-solve, which led to nearly double initial gains in short-term performance compared to the group with access to the basic ChatGPT. When it came to long-term learning, the students with the ChatGPT-4 tutor did not see the same drop in scores as the students with basic ChatGPT; they performed similarly to the control group. Notably, the researchers tested the students after only one practice session with AI. Additionally, the prompting given to the ChatGPT-4 tutor was fairly simple, yet it eliminated the later skill gap between the student groups. Given these factors, as well as the size of the increase in short-term performance, it is worth considering whether ongoing exposure to a better-trained AI tutor could boost long-term learning.

Recent research also shows promise for expanding how AI can support learning. For instance, one study used AI as a "peer" to help students address physics misconceptions,[76] while another provided real-time feedback on group collaboration.[77] Whether AI ultimately supports or undermines learning will depend on how it is designed, implemented, and used in practice — ultimately, it is up to education leaders, educators, ed tech developers, and researchers to find that path forward.

SIDEBAR 2

AI Possibilities for Teachers

Productive struggle is not just for students; it's a fundamental part of how all humans learn and grow, including teachers. Whether developing instructional plans, responding to student needs, or making pedagogical decisions, teachers engage in rich cognitive work that helps refine their practice over time. These challenges are not simply inefficiencies — they are often critical opportunities for teachers to build professional expertise and deepen relationships with students. Teaching is a learning profession, where adults are also learners who continuously improve their craft.

Yet, it is important to acknowledge that the teaching profession, as it currently stands, is unsustainable. Among public school teachers who were teaching in the 2020-21 school year, 16% moved to a different school or left the teaching profession, with even higher rates for those who work at schools serving a large percentage of students from low-income households.[78] Educators face a barrage of demands, and AI has enormous potential to remove unnecessary friction so that teachers can free up time for deep, relational, and intellectually engaging parts of their job.

However, educators must tread carefully. While educators have long used curriculum materials, AI tools are different in an important way: they deliver instant, customized outputs that have the potential to *replace* rather than *support* teacher thinking. If overused or misapplied, this can lead to less engagement with the learning goals, standards, or key ideas that teachers usually consider when creating or adjusting a lesson. Similarly, while AI-generated feedback may be efficient, overreliance may miss the relational nuances that come with authentic teacher-student relationships that are vital for trust, motivation, and growth.

The opportunity is neither to entirely embrace nor reject AI, but to use it wisely: offload tasks that drain capacity without enriching practice, while preserving and amplifying the kinds of productive struggle that lead to professional growth. In doing so, teachers can create space for a more sustainable, human-centered vision for the profession.

Beyond Cognition

THE HUMAN SIDE OF LEARNING

Although the focus of this report is on AI's impact on productive struggle in the academic context, AI also plays a role in the more human side of learning, specifically social development and creativity.

SOCIAL DEVELOPMENT AND RELATIONSHIPS

Schools are more than a place of academic learning; they also help students develop valuable social skills. AI has the opportunity to support competencies such as "self-awareness, empathy, and collaboration."[79] Examples of potential use cases include monitoring facial expressions or tone of voice for real-time support and tailored interventions; generating personalized prompts or journaling exercises; or creating "synthetic personas or 'characters' that expose educators and students to diverse perspectives, fostering empathy and cultural awareness."[80] AI has also shown potential as a therapy tool,[81] which could open up new capabilities in working with students on social development skills.

However, there is the potential for serious risks associated with AI and social development. Some of the risks are related to reduced human interactions and isolation, particularly with extended use.[82] Another potential risk is that some AI chatbots may be overly agreeable, potentially reinforcing a user's thoughts — including risky or dangerous behaviors — instead of challenging them.[83] Finally, there are risks related to the AI models themselves given the potential for underlying biases in the data, which may have a greater impact on minority student groups.[84]

What Early Research Suggests

Although there are general concerns about AI's impact on social development and relationship-building, there is limited research to fully understand the risks. For instance, surveys of faculty and undergraduates surfaced general concern that AI could reduce face-to-face interactions between faculty and students or between peers.[85] Similarly, researchers have concerns that although AI can give students some opportunities to practice social interactions, AI will never be able to fully replicate the real world and students need additional opportunities to practice in those conditions.[86]

Additional research in this area is important because there may be meaningful individual differences that influence the way AI impacts social development. For example, a study of ChatGPT users found that generally, there is little emotional engagement with ChatGPT, but for some individuals, higher daily usage is related to higher levels of "loneliness dependence, and problematic use, and lower socialization."[87] Conversely, while one study found that chatbots can reduce suicidal ideation in some users,[88] other cases, highlighted in recent lawsuits, point to potential harm for other users,[89] particularly those under age 18.[90]

"More broadly, schools are a place where socialization happens and learning about self and others through interactions with others. There's an opportunity to think about how to build in acknowledging that and how we can leverage AI in ways that facilitate that implicit thing that's happening."[91] **—Janis Whitlock, Cornell University**

CREATIVITY

Creativity involves combining ideas in ways that generate "novel value, use, or meaning" for others.[92] In some ways, AI can be supportive of creativity, either as a brainstorming partner or by allowing users to create a range of products from visuals to music to apps much more easily.

What Early Research Suggests

The early research on creativity is mixed. To date, studies suggest that while AI can help with brainstorming and idea generation, AI can also make those tasks less fulfilling.

In one study, undergraduates engaged in a creative brainstorming task without and then with ChatGPT.[93] Although AI could support students' divergent thinking (i.e., generating multiple ideas) and students indicated that the technology was helpful in their brainstorming, some students also indicated that they would have preferred not to use AI.[94] Similarly, in a study of undergraduates participating in a creative writing task, the participants who used ChatGPT reported that the task required less effort but was also significantly less enjoyable and less valued.[95] Students in the control group (i.e., who did not use AI) did not have changes in their levels of enjoyment or task value.[96]

"When we outsource the parts of programming that used to demand our complete focus and creativity, do we also outsource the opportunity for satisfaction? Can we find the same fulfillment in prompt engineering that we once found in problem-solving through code?"[97] **—Matheus Lima, Terrible Software**

There are also concerns that AI may narrow creative output more broadly, even as it helps boost individual creativity for some. For instance, in the undergraduate creative brainstorming task mentioned above, some students noted feeling constrained by the AI suggestions.[98] In a separate study asking participants to create short stories where some participants received story ideas from AI, raters evaluated the AI stories as higher quality: deeming them "more creative, better written, and more enjoyable, especially among less creative writers."[99] However, the stories based on GenAI ideas were more similar to one another, suggesting that AI has the potential to narrow novel content as a whole.[100]

Recommendations

AI's growing role in education raises high-stakes questions. Not just about access, efficiency, or proficiency, but also about what kinds of learning are valued, what kinds of thinkers the educational systems are designed to cultivate, and what responsibilities are distributed across institutions in shaping the conditions for both. This moment demands more than reactions; it calls for recalibration.

The seven interdependent recommendations that follow do not offer quick wins or tidy resolutions. Instead, they point toward the kind of slow, steady work that real progress requires: clarifying what matters, aligning systems accordingly, and advancing research while remaining grounded in what is best for student learning and development. Across the public education sector, the decisions that educators, developers, funders, and policymakers make now will ripple forward for decades.

1. Reimagine and redefine what students need to know and become.

As AI automates more tasks and reshapes the labor market, the profile of student readiness, and therefore the objectives of the educational system, must evolve. Knowledge-building remains necessary, but it is no longer enough. **It is increasingly important to cultivate students' capacity to make meaning, weigh evidence, sustain effort, and exercise critical discernment amid complexity**.

Educational goals should shift to reflect this broader vision. This means intentionally embedding motivation, metacognition, and adaptability into the fabric of learning experiences, not treating them as add-ons. It also means articulating which foundational skills still require deep fluency and which may be responsibly supported by tools without compromising developmental integrity.

These shifts also raise new questions not just about what students learn, but about *where* and *how* that learning happens. As students may use AI beyond the classroom, for instructional and entertainment purposes, educators and system leaders must consider how learning is distributed across settings. As a result, schools may need to redesign not only instructional goals, but also time, supervision, family engagement, and the boundaries of learning.

Funders and policy leaders can support this shift by investing in learning standards and curriculum frameworks that blend cognitive science with real-world applicability. The aim is not to chase novelty, but to ensure students are equipped for the kinds of challenges AI cannot solve for them. As AI takes on more of the routine, the passable bar for what humans contribute will rise. The ability to distinguish among mediocre, good, and truly exceptional, to know when to accept an answer and when to challenge it, will define the value of human judgment. In that future, discernment is not just an academic exercise, it is an essential differentiator.

2. Build coherent systems that align capacity and technology to learning.

AI-related investments will not yield meaningful returns unless the broader educational ecosystem, including how to support teachers and measure learning, shifts in tandem. **These shifts require a coordinated approach across professional development, infrastructure, curriculum, and assessment**.

The ideal approach is likely not a one-off AI workshop or a downloadable resource. It is cultivated through coherent, embedded supports that are part of the everyday fabric of teaching and leadership. Just as schools have come to see internet access and school culture as essential infrastructure, thoughtful use of AI should become part of a school's ethos, not a separate initiative. The gap between what students are doing now and what schools are equipped to address is widening. Closing that gap requires more than awareness; it demands reorientation starting at a class assignment level.

Getting there will take time, and the current landscape is far from that ideal. As of fall 2024, only about half of districts have provided any training to teachers about GenAI, and many of those who did provide training adopted a do-it-yourself approach.[101] In this context, it makes sense to pursue multiple approaches, supporting targeted programs that build AI literacy while also investing in the long-term work of integrating AI into instructional models, coaching structures, and decision-making routines.

Philanthropic funders and system leaders can accelerate progress by investing in organizations that already provide high-quality, embedded support to schools. These partners are well-positioned to help educators and leaders navigate the fast-moving terrain of AI without losing sight of what matters most: using every available tool to help students thrive. When done well, AI capacity-building should not feel like something extra; it should be coherent, integrated, and sustained.

3. Empower educators to redesign assignments for an AI-rich world.

Students are already using AI tools in and outside of the classroom,[102] often without guidance. While early policy responses focus on preventing cheating, the deeper concern is that many instructional practices have not yet adapted to ensure students still engage in meaningful cognitive work. As a result, students may substitute AI for the very critical thinking process that assignments are meant to develop. **Educators need immediate, practical tools to recalibrate what learning looks like in this new context. With thoughtful adjustments, educators can reclaim and even deepen the role of productive struggle, helping students learn not just content but also discernment, iteration, and independent thinking**.

This redesign process could include a shift away from the final product, and instead prioritize tasks that require students to visibly demonstrate their thinking process. Educators may encourage initial attempts without technology so students can grapple with the core cognitive challenge first. When technology is introduced, educators should make sure its role is additive, clarifying confusion, enhancing creativity, or extending thinking, rather than bypassing key skills. Furthermore, some assignments may require meaningful modification such that they focus on what AI cannot easily replicate.

Learning in an AI-rich world still requires effort — though the effort may be less about memorizing and producing, and more about making meaning, evaluating, iterating, and engaging. Empowering educators to redesign assignments with this in mind can protect what matters most: students' development as thoughtful, capable learners.

4. Reinvest in research that reflects the moment.

Decades of research in cognitive science, developmental psychology, and learning theory provide a rich foundation for understanding how students learn and develop. That body of knowledge should be the starting point for how education responds to emerging technologies. At the same time, research is not static, and the rise of AI introduces a wave of new questions that existing studies were never designed to answer.

While this report draws on a growing body of research about AI's role in learning, the field remains uneven, with the voice of educators underrepresented and a need for additional research focused on elementary and secondary students, instead of postsecondary students. Many of the most pressing questions, such as how AI changes student motivation, alters classroom dynamics, or reshapes the role of productive struggle, are not yet fully understood. Recent cuts to federal education research have made it even harder to fill those gaps. Furthermore, the rapid pace of technology advancements complicates the traditional research process; factors like which AI model is used, how it is prompted, and the context with which students interact with it can all shape outcomes.

What is needed is a new wave of interdisciplinary inquiry, bringing together cognitive scientists, developmental psychologists, educators, and technologists, to study how students actually experience AI in real classrooms. These questions will not be answered in lab settings alone. They require research-practice partnerships that connect districts, developers, and academics in sustained, reciprocal ways.

At the same time, the research process itself must evolve. Traditional timelines and funding models are often too slow and siloed to keep up with technological change. AI can be part of the solution here as well, used to supercharge data collection, analysis, and hypothesis testing. Especially in an era of constrained budgets, making the research enterprise more nimble, iterative, and applied will be critical.

5. Reorient measurement toward learning, not just use.

Ed tech tools are often measured by how widely they are used via output measures such as number of users or session length, yet developers and district leaders should resist the gravitational pull toward solely metrics of convenience. **The more important outcomes are slower, harder to measure, and far more consequential: Did this tool deepen engagement? Did it help students transfer knowledge, not just complete tasks? Did it strengthen a teacher's ability to differentiate instruction or foster classroom connection?**

Funders, policymakers, and educational leaders can help shift the incentives by requiring evidence of developmental impact, not just scale. They can also support the creation of common frameworks for evaluating learning tools, grounded in the science of what helps students grow.

6. Develop benchmarks that reflect how students learn.

Most AI benchmarks today measure technical performance, such as how well a system scores on standardized tests or solves complex problems. While these metrics reveal what AI can do, they say little about how it supports students' development. In education, success is not just about correctness; it is about whether tools deepen thinking, promote engagement, and strengthen student-teacher and peer relationships.

There is a particular need for benchmarks that capture motivation and challenge. Few tools are assessed for how well they keep students in the zone of proximal development, where tasks are difficult enough to require effort, but not so hard that students disengage. That balance is crucial for learning, yet it is absent in current benchmarks and evaluation frameworks.

Developers and researchers have the opportunity to work together to create measures that reflect the cognitive, motivational, and relational conditions of real learning. Tools should be assessed not only for accuracy but also for how they shape persistence, curiosity, and long-term understanding across diverse students, including those with learning differences or those learning in multiple languages. Funders can accelerate this shift by prioritizing products that align with the science of learning, not just the speed and power of computation.

7. Center learning science in product design.

AI tools built for education should reflect how students actually learn, including students with learning differences and multilingual students. That means deeply embedding principles from cognitive science, developmental psychology, Universal Design for Learning, and the science of motivation and productive struggle into the design from the outset. Tools built with variability in mind from the start often lead to stronger learning environments for all students, not just those with formal support needs.[103]

Rather than maximizing scale or marketability, educational technology needs to prioritize developmental integrity. One promising model is the use of "red teams" to rigorously test tools. Red teaming should not just be done for security, but also for developmental shortcuts, like over-scaffolding, bypassing effort, or undermining agency. These checks can surface unintended consequences before products reach students.

Good intentions are not enough. Policies and philanthropic investment should create market incentives for product teams to prioritize learning outcomes, not just technical performance or market share. Developers who treat learning science as a core design constraint, rather than a marketing flourish, will be better positioned to build tools that truly benefit students and earn trust in schools.

Conclusion

The scene in Ms. Lopez's class — where Mateo, Jada, and Laila each navigated the same AI tool in strikingly different ways — captures the complexity of this moment. They remind us that technology does not act on students in uniform ways. It interacts with who they are, what they know, how they are supported, and what they are asked to do. The challenge ahead is not simply whether and when AI should be in classrooms, but how its use will shape student effort, identity, and opportunity over time.

As AI capabilities accelerate, education cannot afford to remain reactive. Students cannot wait for a district's AI policy and educators' professional development to slowly catch up. In this moment of flux, the absence of intentional design risks normalizing cognitive offloading as the new norm. The stakes are too high, and the window for intentional design is already narrowing. This is a shared responsibility among educators, developers, funders, and policymakers to ensure that the tools shaping tomorrow's learning reflect what research, experience, and students themselves tell us matters most.

The work ahead lies in making careful distinctions: between scaffolding and shortcut, engagement and distraction, support and substitution. What is needed is not rigid lines but sharper awareness, an ongoing discernment of when AI is extending learning and when AI may be quietly replacing learning. The future of education will not be written by algorithms, but by the values, decisions, and collective courage to use them wisely in ways that expand opportunity and support every student's potential. ✦

Endnotes

1 "NAEP Report Card: Mathematics, Grade 4," The Nation's Report Card, 2024, https://www.nationsreportcard.gov/reports/mathematics/2024/g4_8/?grade=4; "NAEP Report Card: Mathematics, Grade 8," The Nation's Report Card, 2024, https://www.nationsreportcard.gov/reports/mathematics/2024/g4_8/?grade=8; "NAEP Report Card: Reading, Grade 4," The Nation's Report Card, 2024, https://www.nationsreportcard.gov/reports/reading/2024/g4_8/?grade=4; "NAEP Report Card: Reading, Grade 8," The Nation's Report Card, 2024, https://www.nationsreportcard.gov/reports/reading/2024/g4_8/?grade=8.

2 Dan Goldhaber, Thomas J. Kane, Andrew McEachin, and Emily Morton, *A Comprehensive Picture of Achievement Across the COVID-19 Pandemic Years: Examining Variation in Test Levels and Growth Across Districts, Schools, Grades, and Students*, Working Paper No. 266-0522 (CALDER, 2022), https://caldercenter.org/publications/comprehensive-picture-achievement-across-covid-19-pandemic-years-examining-variation, 10–11.

3 States were set to finalize their plans for Elementary and Secondary School Emergency Relief funding by September 2024 and spend the funds by January 2025. Some states received extensions through March 2026; however, the U.S. Department of Education has sought for spending to be complete by May 2025. Juan Perez Jr., "Education Department Halts Final Payouts of Federal Pandemic Relief Funds," Politico, March 28, 2025, https://www.politico.com/news/2025/03/28/education-department-halts-final-payouts-of-federal-pandemic-relief-funds-00258985; Mark Lieberman, "Billions for Schools Are in Limbo as Trump Admin. Denies State Funding Requests," *Education Week*, May 12, 2025, sec. Policy & Politics, Education Funding, https://www.edweek.org/policy-politics/billions-for-schools-are-in-limbo-as-trump-admin-denies-state-funding-requests/2025/05.

4 Krista Kaput, Carrie Hahnel, and Biko McMillan, "How Student Enrollment Declines Are Affecting Education Budgets, Explained in 10 Figures," Bellwether, September 11, 2024, https://bellwether.org/publications/how-student-enrollment-declines-are-affecting-education-budgets/.

5 Jennifer O'Neal Schiess and Bonnie O'Keefe, "Leading Indicator: State Education Finance Issue Three," Bellwether, May 5, 2025, https://bellwether.org/state-education-finance/leading-indicator-state-education-finance-issue-three/.

6 Amy Chen Kulesa, Michelle Croft, Marisa Mission, Brian Robinson, Mary K. Wells, Andrew J. Rotherham, and John Bailey, *Learning Systems: The Landscape of Artificial Intelligence in K-12 Education* (Bellwether, updated May 2025), https://bellwether.org/publications/learning-systems/, 12.

7 Amy Chen Kulesa, Michelle Croft, Brian Robinson, Mary K. Wells, Andrew J. Rotherham, and John Bailey, *Learning Systems: Artificial Intelligence Use Cases* (Bellwether, September 2024), https://bellwether.org/publications/learning-systems/.

8 Ethan Mollick, "A New Generation of AIs: Claude 3.7 and Grok 3," One Useful Thing, February 24, 2025, https://www.oneusefulthing.org/p/a-new-generation-of-ais-claude-37. See for examples of advanced reasoning models: "Grok 3 Beta — the Age of Reasoning Agents," xAI, February 19, 2025, https://x.ai/blog/grok-3; "Claude 3.7 Sonnet and Claude Code," Anthropic, February 24, 2025, https://www.anthropic.com/news/claude-3-7-sonnet; Koray Kavukcuoglu, "Gemini 2.5: Our Most Intelligent AI Model," Google, *The Keyword* (blog), March 25, 2025, https://blog.google/technology/google-deepmind/gemini-model-thinking-updates-march-2025/; "Introducing OpenAI o3 and o4-mini," OpenAI, April 16, 2025, https://openai.com/index/introducing-o3-and-o4-mini/.

9 See for example: Juraj Gottweis and Vivek Natarajan, "Accelerating Scientific Breakthroughs With an AI Co-Scientist," Google, *Research* (blog), February 19, 2025, https://research.google/blog/accelerating-scientific-breakthroughs-with-an-ai-co-scientist/.

10 See for examples: Dave Citron, "Try Deep Research and Our New Experimental Model in Gemini, Your AI Assistant," Google, *The Keyword* (blog), December 11, 2024, https://blog.google/products/gemini/google-gemini-deep-research/; "Introducing Deep Research," OpenAI, February 2, 2025, https://openai.com/index/introducing-deep-research/; "Introducing Perplexity Deep Research," Perplexity, February 14, 2025, https://www.perplexity.ai/hub/blog/introducing-perplexity-deep-research.

11 See results from the GPQA Diamond test on "AI Benchmarking Hub," Epoch AI, updated May 9, 2025, https://epoch.ai/data/ai-benchmarking-dashboard.

12 For example, OpenAI's o3 and o4-mini models incorporate computer vision such that the models "don't just see an image — they think with it." "Introducing OpenAI o3 and o4-mini," OpenAI, April 16, 2025, https://openai.com/index/introducing-o3-and-o4-mini/.

13 Claire Zau, "DeepSeek R-1 Explained," *GSV: AI & Education*, Substack, January 28, 2025, https://aieducation.substack.com/p/deepseek-r-1-explained.

14 Amy Chen Kulesa, Michelle Croft, Marisa Mission, Brian Robinson, Mary K. Wells, Andrew J. Rotherham, and John Bailey, *Learning Systems: Opportunities and Challenges of Artificial Intelligence-Enhanced Education* (Bellwether, updated May 2025), https://bellwether.org/publications/learning-systems/.

15 Robert A. Bjork and Marcia C. Linn, "The Science of Learning and the Learning of Science," *APS Observer* 19, no. 3 (March 2006).

16 Saul McLeod, "Vygotsky's Zone of Proximal Development," Simply Psychology, March 26, 2025, https://www.simplypsychology.org/zone-of-proximal-development.html, citing Lev S. Vygotsky, *Mind in Society: The Development of Higher-Psychological Processes* (Cambridge, MA: Harvard University Press, 1978).

17 Jamaal Rashad Young, Danielle Bevan, and Miriam Sanders, "How Productive Is the Productive Struggle? Lessons Learned from a Scoping Review," *The International Journal of Education in Mathematics, Science, and Technology* 12, no. 2 (2024): 470–95, https://files.eric.ed.gov/fulltext/EJ1413403.pdf, 471.

18 Ibid., 474; Jason M. Lodge, Gregor Kennedy, Lori Lockyer, Amael Arguel, and Mariya Pachman, "Understanding Difficulties and Resulting Confusion in Learning: An Integrative Review," *Frontiers in Education* 3 (June 28, 2018), https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2018.00049/full.

19 Young et al., "How Productive Is the Productive Struggle? Lessons Learned from a Scoping Review," 474; Lodge et al., "Understanding Difficulties and Resulting Confusion in Learning: An Integrative Review."

20 Don W. Edgar, "Learning Theories and Historical Events Affecting Instructional Design in Education: Recitation Literacy Toward Extraction Literacy Practices," *SAGE Open* 2, no. 4 (October 1, 2012): 1–9, https://journals.sagepub.com/doi/pdf/10.1177/2158244012462707, 2–5.

21 Alan Baddeley, "Working Memory: Looking Back and Looking Forward," *Nature* 4 (October 2023): 829–39, https://home.csulb.edu/~cwallis/382/readings/482/baddeley.pdf. It should be noted that there are other mechanisms influencing the flow from sensory input to long-term memory, such as attention and prior knowledge. However, this model serves as a useful example to understand the basic process.

22 For instance, psychologists have proposed an "episodic buffer" as part of the working memory system. The buffer is a temporary storage space that integrates information from different sources, such as visual, auditory, and potentially smell and taste, along with data from long-term memory. Alan Baddeley, "Working Memory," *Current Biology* 20, no. 4 (February 23, 2010): R136–40, https://www.cell.com/current-biology/abstract/S0960-9822(09)02133-2.

23 Baddeley, "Working Memory"; National Academies of Sciences, Engineering, and Medicine, "Chapter 4," in *How People Learn II: Learners, Contexts, and Cultures* (Washington, DC: The National Academies Press, 2018), https://doi.org/10.17226/24783, 74–75, 80; Dan Williams, "The Importance of Cognitive Load Theory," Society for Education & Training, https://set.et-foundation.co.uk/resources/the-importance-of-cognitive-load-theory.

24 Williams, "The Importance of Cognitive Load Theory."

25 Patricia Ann DeWinstanley and Elizabeth Ligon Bjork, "Processing Strategies and the Generation Effect: Implications for Making a Better Reader," *Memory & Cognition* 32, no. 6 (2004): 945–55, https://bjorklab.psych.ucla.edu/wp-content/uploads/sites/13/2016/07/DeWinstanley_EBjork_2004.pdf; Kathleen B. McDermott and Henry L. Roediger III, "Memory (Encoding, Storage, Retrieval)," Noba, 2025, https://nobaproject.com/modules/memory-encoding-storage-retrieval; National Academies of Sciences, Engineering, and Medicine, "Chapter 4," in *How People Learn II: Learners, Contexts, and Cultures*, 75–81.

26 DeWinstanley and Bjork, "Processing Strategies and the Generation Effect: Implications for Making a Better Reader"; McDermott and Roediger III, "Memory (Encoding, Storage, Retrieval)"; National Academies of Sciences, Engineering, and Medicine. 2018. *How People Learn II: Learners, Contexts, and Cultures.* National Academies of Sciences, Engineering, and Medicine, "Chapter 4," in How People Learn II: Learners, Contexts, and Cultures, 75–81.

27 See Elizabeth A. Styles, *The Psychology of Attention* (Bucks, UK: Hove and New York, 2005), https://www.psicopolis.com/boxtesti/psicattenz.pdf.

28 See Elizabeth A. Styles, "3. Selective Report and Interference Effects in Visual Attention," in *The Psychology of Attention*, 26–46.

29 National Academies of Sciences, Engineering, and Medicine, "Chapter 4," in *How People Learn II: Learners, Contexts, and Cultures*, 71.

30 Ibid.

31 Ibid. See also Faria Sana, Tina Weston, and Nicholas J. Cepeda, "Laptop Multitasking Hinders Classroom Learning for Both Users and Nearby Peers," *Computers & Education* 62 (March 1, 2013): 24–31, https://www.sciencedirect.com/science/article/pii/S0360131512002254.

32 See Vicki Trowler, "Student Engagement Literature Review," *The Higher Education Academy*, November 2010, https://www.improvingthestudentexperience.com/wp-content/uploads/library/UG_documents/StudentEngagementLiteratureReview_2010.pdf.

33 Mihaly Csikszentmihalyi, "Introduction," in *Flow: The Psychology of Optimal Experience* (Harper, 1990), 1–8, https://www.researchgate.net/publication/224927532_Flow_The_Psychology_of_Optimal_Experience.

34 Amy McDermott, "When Kids Focus on Challenges in Short Spurts, They Build 'Cognitive Endurance,'" *Journal Club* (blog) (Proceedings of the National Academy of Sciences, January 10, 2025), https://www.pnas.org/post/journal-club/kids-focus-challenges-short-spurts-they-build-cognitive-endurance.

35 Kenneth W. K. Lo, Grace Ngai, Stephen C. F. Chan, and Kam-por Kwan, "How Students' Motivation and Learning Experience Affect Their Service-Learning Outcomes: A Structural Equation Modeling Analysis," *Frontiers in Psychology* 13 (April 18, 2022), https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.825902/full.

36 Richard M. Ryan and Edward L. Deci, "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," *American Psychologist* 55, no. 1 (January 2000): 68–78, https://selfdeterminationtheory.org/SDT/documents/2000_RyanDeci_SDT.pdf.

37 Lo et al., "How Students' Motivation and Learning Experience Affect Their Service-Learning Outcomes: A Structural Equation Modeling Analysis."

38 Ibid.

39 For a more comprehensive overview of expectancy-value theory, see Hasan Bircan and Semra Sungur, "The Role of Motivation and Cognitive Engagement in Science Achievement," *Science Education International* 27, no. 4 (2016): 509–29, https://files.eric.ed.gov/fulltext/EJ1131144.pdf, 509–13.

40 Ryan and Deci, "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," 70.

41 Bircan and Sungur, "The Role of Motivation and Cognitive Engagement in Science Achievement," 511–14.

42 "Growth mindset" is the most popular term, but stems from attribution theory in motivational research: Carol S. Dweck and Ellen L. Leggett, "A Social-Cognitive Approach to Motivation and Personality," Psychological Review 95, no. 2 (1988): 256–73, https://mathedseminar.pbworks.com/f/Dweck+&+Leggett+(1988)+A+social-cognitive+approach+to+motivation+and+personality.pdf, 268–269. See also David S. Yeager and Carol S. Dweck, "What Can Be Learned from Growth Mindset Controversies?," *The American Psychologist* 75, no. 9 (December 2020): 1269–84, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8299535/.

43 Young et al., "How Productive Is the Productive Struggle? Lessons Learned from a Scoping Review," 472.

44 Jennifer G. Cromley and Andrea J. Kunze, "Metacognition in Education: Translational Research," *Translational Issues in Psychological Science* 6, no. 1 (2020): 15–20, https://experts.illinois.edu/en/publications/metacognition-in-education-translational-research; "Metacognition," *MIT Teaching + Learning Lab* (blog), https://tll.mit.edu/teaching-resources/how-people-learn/metacognition/.

45 Gregory Schraw and David Moshman, "Metacognitive Theories," *Educational Psychology Review* 7, no. 4 (December 3, 1995): 351–71, https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1040&context=edpsychpapers, 351.

46 Emily Lai, *Metacognition: A Literature Review Research Report* (Pearson, January 1, 2011), https://www.academia.edu/64842513/Metacognition_A_Literature_Review_Research_Report, 5, citing Deanna Kuhn and David Dean, Jr., "Metacognition: A Bridge Between Cognitive Psychology and Educational Practice," *Theory Into Practice* 43, no. 4 (November 1, 2004): 268–73, https://www.tandfonline.com/doi/abs/10.1207/s15430421tip4304_4.

47 Philip H. Winne and Roger Azevedo, "Metacognition and Self-Regulated Learning," in *The Cambridge Handbook of the Learning Sciences*, ed. R. Keith Sawyer, 3rd ed., Cambridge Handbooks in Psychology (Cambridge: Cambridge University Press, 2022), 93–113, https://doi.org/10.1017/9781108888295.007.

48 Young et al., "How Productive Is the Productive Struggle? Lessons Learned from a Scoping Review."

49 Ibid.

50 "Using Retrieval Practice to Increase Student Learning," Center for Teaching and Learning, https://ctl.wustl.edu/resources/using-retrieval-practice-to-increase-student-learning/.

51 Haley A. Vlach and Catherine M. Sandhofer, "Distributing Learning Over Time: The Spacing Effect in Children's Acquisition and Generalization of Science Concepts," *Child Development* 83, no. 4 (July 2012): 1137–44, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3399982/.

52 *Cheat or Be Cheated? What We Know About Academic Integrity in Middle & High Schools & What We Can Do About It* (Challenge Success, 2012), https://challengesuccess.org/wp-content/uploads/2021/04/ChallengeSuccess-AcademicIntegrity-WhitePaper.pdf; Vignesh Ramachandran, Vignesh Ramachandran, "Young People Cheating a Bit Less These Days, Report Finds," NBC News, November 26, 2012, http://www.nbcnews.com/news/us-news/young-people-cheating-bit-less-these-days-report-finds-flna1C7265474.

53 See Bree Dusseault, "New State AI Policies Released: Signs Point to Inconsistency and Fragmentation – Center on Reinventing Public Education," CRPE, March 5, 2024, https://crpe.org/new-state-ai-policies-released-inconsistency-and-fragmentation/; Bree Dusseault and Justin Lee, "Study: How Districts Are Responding to AI & What It Means for the New School Year," The 74, September 10, 2023, https://www.the74million.org/article/study-how-districts-are-responding-to-ai-what-it-means-for-the-new-school-year/; Michelle Croft, Nora Weber, and Kelly Robson Foster, *Surveying School System Leaders: What They Are Saying About Artificial Intelligence (So Far)* (Bellwether, February 13, 2025), https://bellwether.org/publications/surveying-artificial-intelligence-and-schools/.

54 Victor R. Lee, Denise Pope, Sarah Miles, and Rosalía C. Zárate, "Cheating in the Age of Generative AI: A High School Survey Study of Cheating Behaviors before and after the Release of ChatGPT," *Computers and Education: Artificial Intelligence* 7 (December 1, 2024), https://www.sciencedirect.com/science/article/pii/S2666920X24000560.

55 Arianna Prothero, "New Data Reveal How Many Students Are Using AI to Cheat," *Education Week*, April 25, 2024, sec. Technology, Artificial Intelligence, https://www.edweek.org/technology/new-data-reveal-how-many-students-are-using-ai-to-cheat/2024/04.

56 Amy Chen Kulesa, Michelle Croft, Marisa Mission, Brian Robinson, Mary K. Wells, Andrew J. Rotherham, and John Bailey, *Learning Systems: Opportunities and Challenges of Artificial Intelligence-Enhanced Education* (Bellwether, updated May 2025), https://bellwether.org/publications/learning-systems/.

57 Olivia Sidoti, Eugenie Park, and Jeffrey Gottfried, "About a Quarter of U.S. Teens Have Used ChatGPT for Schoolwork – Double the Share in 2023," *Pew Research Center* (blog), January 15, 2025, https://www.pewresearch.org/short-reads/2025/01/15/about-a-quarter-of-us-teens-have-used-chatgpt-for-schoolwork-double-the-share-in-2023/.

58 Ibid.

59 *The State of AI in Education 2025: Key Findings from a National Survey* (Carnegie Learning, 2025), https://discover.carnegielearning.com/hubfs/PDFs/Whitepaper%20and%20Guide%20PDFs/2025-AI-in-Ed-Report.pdf?hsLang=en; Caitlynn Peetz, "What Teacher PD on AI Should Look Like. Some Early Models Are Emerging," *Education Week*, December 9, 2024, sec. Technology, Artificial Intelligence, https://www.edweek.org/technology/what-teacher-pd-on-ai-should-look-like-some-early-models-are-emerging/2024/12.

60 Croft et al., *Surveying School System Leaders: What They Are Saying About Artificial Intelligence (So Far)*.

61 Ibid.

62 "Should I Use ChatGPT to Write My Essays?," *Harvard Summer School* (blog), September 6, 2023, https://summer.harvard.edu/blog/should-i-use-chatgpt-to-write-my-essays/; Kristen Purcell, Judy Buchanan, and Linda Friedrich, "Part II: How Much, and What, Do Today's Middle and High School Students Write?," *Pew Research Center* (blog), July 16, 2013, https://www.pewresearch.org/internet/2013/07/16/part-ii-how-much-and-what-do-todays-middle-and-high-school-students-write/.

63 Ibid.

64 Jin Wang and Wenxiang Fan, "The Effect of ChatGPT on Students' Learning Performance, Learning Perception, and Higher-Order Thinking: Insights from a Meta-Analysis," *Humanities and Social Sciences Communications* 12 (May 6, 2025), https://www.nature.com/articles/s41599-025-04787-y.

65 It is important to note that different large language models have different capabilities, which can affect how useful they are for a particular task. In addition, many studies do not include specific information about the prompts they used, even though prompts play a key role in producing quality results.

66 Matthias Stadler, Maria Bannert, and Michael Sailer, "Cognitive Ease at a Cost: LLMs Reduce Mental Effort but Compromise Depth in Student Scientific Inquiry," *Computers in Human Behavior* 160 (November 1, 2024): 108386, https://www.sciencedirect.com/science/article/pii/S0747563224002541.

67 Ibid.

68 Ibid.

69 Hamsa Bastani, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman, "Generative AI Can Harm Learning," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, July 15, 2024), https://papers.ssrn.com/abstract=4895486.

70 Ibid.

71 Yizhou Fan, Luzhen Tang, Huixiao Le, Kejie Shen, Shufang Tan, Yueying Zhao, Yuan Shen, Xinyu Li, and Dragan Gašević, "Beware of Metacognitive Laziness: Effects of Generative Artificial Intelligence on Learning Motivation, Processes, and Performance," *British Journal of Educational Technology* 56, no. 2 (2025): 489–530, https://doi.org/10.1111/bjet.13544.

72 Ibid.

73 Kunal Handa, Drew Bent, Alex Tamkin, Miles McCain, Esin Durmus, Michael Stern, Mike Schiraldi, Saffron Huang, Stuart Ritchie, Steven Syverud, Kamya Jagadish, Margaret Vo, Matt Bell, and Deep Ganguli, "Anthropic Education Report: How University Students Use Claude," Anthropic, April 8, 2025, https://www.anthropic.com/news/anthropic-education-report-how-university-students-use-claude.

74 Octavian-Mihai Machidon, "Generative AI and Childhood Education: Lessons from the Smartphone Generation," *AI & SOCIETY*, February 14, 2025, https://link.springer.com/article/10.1007/s00146-025-02196-y; Chunpeng Zhai, Santoso Wibowo, and Lily D. Li, "The Effects of Over-Reliance on AI Dialogue Systems on Students' Cognitive Abilities: A Systematic Review," *Smart Learning Environments* 11, no. 1 (June 18, 2024): 28, https://slejournal.springeropen.com/articles/10.1186/s40561-024-00316-7.

75 Bastani et al., "Generative AI Can Harm Learning."

76 Ruben Weijers, Denton Wu, Hannah Betts, Tamara Jacod, Yuxiang Guan, Vidya Sujaya, Kushal Dev, Toshali Goel, William Delooze, Reihaneh Rabbany, Ying Wu, Jean-François Godbout, and Kellin Pelrine, "From Intuition to Understanding: Using AI Peers to Overcome Physics Misconceptions" (arXiv, April 1, 2025), https://arxiv.org/abs/2504.00408.

77 Sangwon Seo, Bing Han, Rayan E. Harari, Roger D. Dias, Marco A. Zenati, Eduardo Salas, and Vaibhav Unhelkar, "Socratic: Enhancing Human Teamwork via AI-Enabled Coaching" (arXiv, February 24, 2025), https://arxiv.org/abs/2502.17643.

78 Soheyla Taie and Laurie Lewis, *Teacher Attrition and Mobility Results From the 2021–22 Teacher Follow-up Survey to the National Teacher and Principal Survey* (National Center for Education Statistics, December 2023), https://nces.ed.gov/use-work/resource-library/report/first-look-ed-tab/teacher-attrition-and-mobility-results-2021-22-teacher-follow-survey-national-teacher-and-principal?pubid=2024039.

79 Danah Henriksen, Edwin Creely, Natalie Gruber, and Sean Leahy, "Social-Emotional Learning and Generative AI: A Critical Literature Review and Framework for Teacher Education," *Journal of Teacher Education* 76, no. 3 (May 2025): 312–28, https://research.monash.edu/en/publications/social-emotional-learning-and-generative-ai-a-critical-literature.

80 Henriksen et al., "Social-Emotional Learning and Generative AI: A Critical Literature Review and Framework for Teacher Education."

81 Michael V. Heinz, Daniel M. Mackin, Brianna M. Trudeau, Sukanya Bhattacharya, Yinzhou Wang, Haley A. Banta, Abi D. Jewett, Abigail J. Salzhauer, Tess Z. Griffin, and Nicholas C. Jacobson, "Randomized Trial of a Generative AI Chatbot for Mental Health Treatment," *NEJM AI* 2, no. 4 (March 27, 2025), https://ai.nejm.org/doi/full/10.1056/AIoa2400802; Nabil Saleh Sufyan, Fahmi H. Fadhel, Saleh Safeer Alkhathami, and Jubran Y. A. Mukhadi, "Artificial Intelligence and Social Intelligence: Preliminary Comparison Study between AI Models and Psychologists," *Frontiers in Psychology* 15 (February 2, 2024), https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1353022/full.

82 Henriksen et al., "Social-Emotional Learning and Generative AI: A Critical Literature Review and Framework for Teacher Education"; "Will Artificial Intelligence Make Us Lonelier?," *The Survey Center on American Life* (blog), May 4, 2023, https://www.americansurveycenter.org/newsletter/will-artificial-intelligence-make-us-lonelier/; Cathy Mengying Fang, Auren R. Liu, Valdemar Danry, Eunhae Lee, Samantha W. T. Chan, Pat Pataranutaporn, Pattie Maes, Jason Phang, Michael Lampe, Lama Ahmad, and Sandhini Agarwal, *How AI and Human Behaviors Shape Psychosocial Effects of Chatbot Use: A Longitudinal Randomized Controlled Study* (arXiv, March 21, 2025), https://www.media.mit.edu/publications/how-ai-and-human-behaviors-shape-psychosocial-effects-of-chatbot-use-a-longitudinal-controlled-study/.

83 John Nosta, "AI Is Cognitive Comfort Food," Psychology Today, April 21, 2025, https://www.psychologytoday.com/us/blog/the-digital-self/202504/ai-is-cognitive-comfort-food; Radhika Rajkumar, "OpenAI Recalls GPT-4o Update for Being Too Agreeable," ZDNET, April 30, 2025, https://www.zdnet.com/article/gpt-4o-update-gets-recalled-by-openai-for-being-too-agreeable/.

84 Henriksen et al., "Social-Emotional Learning and Generative AI: A Critical Literature Review and Framework for Teacher Education."

85 Ahmed M. Hasanein and Abu Elnasr E. Sobaih, "Drivers and Consequences of ChatGPT Use in Higher Education: Key Stakeholder Perspectives," *European Journal of Investigation in Health, Psychology and Education* 13, no. 11 (November 9, 2023): 2,599–2,614, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10670526/.

86 Henriksen et al., "Social-Emotional Learning and Generative AI: A Critical Literature Review and Framework for Teacher Education."

87 Fang et al., *How AI and Human Behaviors Shape Psychosocial Effects of Chatbot Use: A Longitudinal Randomized Controlled Study*; "Early Methods for Studying Affective Use and Emotional Well-Being on ChatGPT," OpenAI, March 21, 2025, https://openai.com/index/affective-use-study/.

88 Bethanie Maples, Merve Cerit, Aditya Vishwanath, and Roy Pea, "Loneliness and Suicide Mitigation for Students Using GPT3-Enabled Chatbots," *Npj Mental Health Research* 3, no. 1 (January 22, 2024): 1–6, https://www.nature.com/articles/s44184-023-00047-6.

89 Dan Jasnow, "New Lawsuits Targeting Personalized AI Chatbots Highlight Need for AI Quality Assurance and Safety Standards," *The National Law Review*, January 6, 2025, https://natlawreview.com/article/new-lawsuits-targeting-personalized-ai-chatbots-highlight-need-ai-quality-assurance.

90 AI Risk Assessment Team, "Social AI Companions," Common Sense Media, April 28, 2025, https://www.commonsensemedia.org/ai-ratings/social-ai-companions.

91 Bellwether interview, April 30, 2025.

92 Sabrina Habib, Thomas Vogel, Xiao Anli, and Evelyn Thorne, "How Does Generative Artificial Intelligence Impact Student Creativity?," *Journal of Creativity* 34, no. 1 (April 1, 2024): 100072, https://www.sciencedirect.com/science/article/pii/S2713374523000316, citing: Alice Chirico, Glaveanu, Vlad Petre, Cipresso, Pietro, Riva,Giuseppe, and Andrea and Gaggioli, "Awe Enhances Creative Thinking: An Experimental Study," *Creativity Research Journal* 30, no. 2 (April 3, 2018): 123–31, https://doi.org/10.1080/10400419.2018.1446491; Mihaly Csikszentmihalyi, "On Runco's Problem Finding, Problem Solving, and Creativity," *Creativity Research Journal* 9, no. 2–3 (1996): 267–68, https://www.tandfonline.com/doi/pdf/10.1080/10400419.1996.9651177; J. P. Guilford, "Creativity," *American Psychologist* 5, no. 9 (1950): 444–54, https://psycnet.apa.org/doiLanding?doi=10.1037%2Fh0063487; Joy Paul Guilford, *The Nature of Human Intelligence*. (New York, NY: McGraw-Hill, 1967), https://psycnet.apa.org/record/1967-35015-000; Ellis Paul Torrance, *Creativity*, vol. 13 (Dimensions Publishing Company, 1969), https://files.eric.ed.gov/fulltext/ED078435.pdf; Ellis Paul Torrance, "Torrance Tests of Creative Thinking," *Educational and Psychological Measurement*, 1966, https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft05532-000; Paul Collard and Janet Looney, "Nurturing Creativity in Education," *European Journal of Education* 49, no. 3 (2014): 348–64, https://onlinelibrary.wiley.com/doi/abs/10.1111/ejed.12090.

93 Habib et al., "How Does Generative Artificial Intelligence Impact Student Creativity?"

94 Ibid.

95 Peidong Mei, Deborah N. Brewis, Fortune Nwaiwu, Deshan Sumanathilaka, Fernando Alva-Manchego, and Joanna Demaree-Cotton, "If ChatGPT Can Do It, Where Is My Creativity? Generative AI Boosts Performance but Diminishes Experience in Creative Writing," *Computers in Human Behavior: Artificial Humans* 4 (May 1, 2025): 100–40, https://www.sciencedirect.com/science/article/pii/S2949882125000246.

96 Ibid.

97 Matheus Lima, "The Hidden Cost of AI Coding," *Terrible Software* (blog), April 23, 2025, https://terriblesoftware.org/2025/04/23/the-hidden-cost-of-ai-coding/.

98 Habib et al., "How Does Generative Artificial Intelligence Impact Student Creativity?"

99 Anil R. Doshi and Oliver P. Hauser, "Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content," *Science Advances* 10, no. 28 (July 12, 2024), https://www.science.org/doi/10.1126/sciadv.adn5290.

100 Ibid.

101 Melissa Kay Diliberti, Robin J. Lake, and Steven R. Weiner, *More Districts Are Training Teachers on Artificial Intelligence: Findings from the American School District Panel* (RAND, April 8, 2025), https://www.rand.org/pubs/research_reports/RRA956-31.html.

102 Olivia Sidoti et al., "About a Quarter of U.S. Teens Have Used ChatGPT for Schoolwork – Double the Share in 2023."

103 See "Universal Design for Learning," CAST, https://www.cast.org/what-we-do/universal-design-for-learning/; Matthew James Capp, "The Effectiveness of Universal Design for Learning: A Meta-Analysis of Literature between 2013 and 2016," *International Journal of Inclusive Education* 21, no. 8 (August 3, 2017): 791–807, https://doi.org/10.1080/13603116.2017.1325074.

About the Authors

AMY CHEN KULESA

Amy Chen Kulesa is a senior associate partner at Bellwether in the Strategic Advising practice area. She can be reached at **amy.chenkulesa@bellwether.org**.

MARISA MISSION

Marisa Mission is a senior analyst at Bellwether in the Policy and Evaluation practice area. She can be reached at **marisa.mission@bellwether.org**.

MICHELLE CROFT

Michelle Croft is an associate partner at Bellwether in the Policy and Evaluation practice area. She can be reached at **michelle.croft@bellwether.org**.

MARY K. WELLS

Mary K. Wells is a co-founder and senior partner at Bellwether in the Strategic Advising practice area. She can be reached at **mary@bellwether.org**.

About Bellwether

Bellwether is a national nonprofit that exists to transform education to ensure systemically marginalized young people achieve outcomes that lead to fulfilling lives and flourishing communities. Founded in 2010, we work hand in hand with education leaders and organizations to accelerate their impact, inform and influence policy and program design, and share what we learn along the way. For more, visit **bellwether.org**.

ACKNOWLEDGMENTS

We would like to thank the many experts who gave their time and shared their knowledge with us to inform our work, including John Bailey and all of the experts we have interviewed over the past year. Thank you also to the Bezos Family Foundation, Chan Zuckerberg Initiative, Charter School Growth Fund, and Overdeck Family Foundation for their financial support of this project.

We would also like to thank our Bellwether colleagues Biko McMillan, Emily Shisler, and Melissa Steel King for their input and Janine Sandy for her support. Thank you to Amy Ribock, Kate Neifeld, Andy Jacob, McKenzie Maxson, Zoe Cuddy, Julie Nguyen, Mandy Berman, and Amber Walker for shepherding and disseminating this work, and to Super Copy Editors.

The contributions of these individuals and entities significantly enhanced our work; however, any errors in fact or analysis remain the responsibility of the authors.

Bellwether

FORWARD THINKING. FORWARD MOVING.