Community for Advancing Discovery Research in Education

Generative AI in STEM Teaching: Opportunities and Tradeoffs

JANUARY 2025

Acknowledgments

This brief is a product of the Community for Advancing Discovery Research in Education (CADRE) Committee on AI in STEM Education and Education Research. CADRE connects STEM education researchers funded by the National Science Foundation's Discovery Research PreK-12 (DRK-12) program to share their methods, findings, results, and products with the research and development community and the greater public, to better inform and impact policy, research and education.

Steering Committee:

Christopher J. Harris,
WestEd
Eric Wiebe,
North Carolina State University
Shuchi Grover,
Looking Glass Ventures

Reviewer:

Ebony Terrell Shockley,
University of Maryland College Park

Authors:

Jeremy Forest Price,
Indiana University Indianapolis
Shuchi Grover,
Looking Glass Ventures

Suggested Citation:

Price, J.F., & Grover, S. (2025). *Generative AI in STEM teaching: Opportunities and tradeoffs*. Community for Advancing Discovery Research in Education (CADRE). Education Development Center, Inc.

The activity was funded by the National Science Foundation, Grant # 2100823. Any opinions, findings, and conclusions or recommendations expressed in these materials are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Generative AI was not used to create this brief. Photos from Pexels.com. Credits: Pavel Danilyuk, Vlada Karpovich, Charlotte May, and Max Fischer.

Table of Contents

Foreword

By Christopher J. Harris and Eric Wiebe

Research and development work in artificial intelligence in education (AIED) is wide ranging and rapidly growing in support of all areas of science, technology, engineering, and mathematics (STEM) teaching and learning. With its broad applicability and transformative potential, AIED represents what could arguably be the most fundamentally game-changing technology for education to emerge since the Internet. Building from prior decades of work on artificial intelligence (AI) and AI-based learning and teaching technologies, the recent advances in AIED—especially the leaps in generative AI—are pushing us to reimagine what is possible for STEM teaching and learning. AIED research initiatives are being prioritized for funding and AIED advances are quickly becoming integrated into STEM education. It is transforming how teachers teach and how students learn. It is also transforming how education developers and researchers conduct their work. There is excitement about the promise of AIED as well as growing concern that the breakthroughs in AIED are impacting everyday education practice in ways that may perpetuate long-standing biases and diminish the potential for broad-based improvement in outcomes. As the research and development work continues to push the boundaries of what is possible, it is important to keep the dual promise and peril of AIED in mind.

This brief is the second in a three-part series on AIED. The series' topics address ethical approaches to AI in STEM education research, AI for STEM teaching, and AI for STEM learning. This series is sponsored by the *Community for Advancing Discovery Research in Education* (CADRE), an NSF-funded network for STEM education researchers endeavoring to improve STEM teaching and learning through research, development, and various information-sharing and community-building mechanisms. Researchers in the CADRE network are part of a portfolio of projects funded through the National Science Foundation's (NSF's) *Discovery Research PreK-12* (DRK-12) program. The DRK-12 portfolio is a wide-ranging set of projects that focus on applied research and development to generate innovative research-informed and field-tested tools, products, and approaches that are intended to enhance STEM teaching and learning. Over the past several years, the portfolio has grown to include an increasing number of projects that leverage AIED to achieve their goals related to teaching or learning. It is expected to continue to grow. This series has been inspired by the question, "What are the essential considerations for researchers and developers who are designing, studying, and using AI in K–12 STEM classrooms?" Our hope is that the opportunities and challenges discussed in this series will generate reflection and rich discussion that furthers the ethical and transformative use of AI to achieve positive and wide-reaching impact for STEM educators and learners.

In this second brief, *Generative AI in STEM Teaching: Opportunities and Tradeoffs*, the authors Jeremy Price and Shuchi Grover explore the exciting possibilities and critical challenges of generative AI (GenAI) for STEM teaching. They examine how GenAI may shape STEM classrooms in the near future and identify promising trajectories for integrating GenAI into STEM teaching, including the potential for personalized teaching approaches. This brief delves into the current state of GenAI research and development around five dimensions of STEM teaching:

1. Justice, equity, and inclusion
2. Curricular decision-making and lesson planning
3. Pedagogy and instruction
4. Assessing student work and progress
5. Teacher learning and leadership

Throughout the brief, the authors reaffirm the essential role of teachers and emphasize the need to include teachers in the design, development, and study of GenAI tools that are intended to benefit STEM education. They also highlight areas of critical concern, urging those involved in research and development to prioritize GenAI efforts in addressing the substantive problems of practice central to STEM teaching with an overarching aim to create more inclusive learning environments.

All told, this brief offers a valuable perspective that encourages us to consider the profound potential impact of GenAI for supporting teachers and improving STEM education on a large scale. The emphasis being on *potential*—in which emerging GenAI tools may help teachers navigate the complexities of STEM teaching and strengthen their capacity to deliver high-quality learning experiences that are calibrated to the needs of all students in today's STEM classrooms.

Introduction and Motivation

Since the momentous arrival of ChatGPT, generative artificial intelligence (GenAI) has seen a significant uptake in K–12 classrooms. According to a recent RAND report (Diliberti et al., 2024), educators are increasingly incorporating AI tools into their teaching practices. The hype surrounding GenAI tools, such as MagicSchool; StretchAI, the AI coach for educators by the International Society of Technology in Education (ISTE); and others, has been met with both hope and skepticism about their promise of improving teaching and learning and creating more inclusive learning environments.

Historically, large-scale AI in education (AIED) efforts date back to the 1980s with intelligent tutoring systems (ITSs) aimed at assisting students through personalized learning. Early efforts evolved as algorithms that modeled student learning became more sophisticated; however, the capabilities of such systems for tailoring learning to individual student's needs were limited by the lack of computing power and access to large corpi of data. With the rise of big data and the attendant rise of "learning analytics" and "educational data mining" (EDM)—subfields at the intersection of computer science and education—AI-based tools began to move beyond being cognitive tutors to systems that embraced socio-emotional factors in student learning as well as supporting the teacher by providing in-the-moment insights from digital learning environments. However, the research and development culture, shaped in part by the technical complexity in working with the underlying platforms, usually resulted in minimal teacher input in their development. Additionally, adoption of such technologies was slower than expected—it typically required stakeholder buy-in, planning, expenditure, and teacher preparation.

In contrast, the latest machine learning (ML) and GenAI techniques have produced a slew of foundational large language models (LLMs), such as OpenAI's ChatGPT, Anthropic's Claude, Microsoft's Co-Pilot, and Google's Gemini, and they are poised to empower and support teachers in hitherto unimaginable ways, promising a more general-purpose and democratic approach to AI in STEM education. This is largely due to the fact that pre-GenAI research-based AIED tools represented what is generally viewed as "narrow AI"—AI that is designed to perform specific tasks or a limited range of tasks and is focused on a single domain or application area (e.g., detecting a student's struggle with specific mathematics tasks in a middle school curriculum). On the other hand, "general-purpose AI" has the capabilities of recent LLMs that appear to match human-level intelligence across a wide range of cognitive tasks, and which can draw on and apply knowledge across multiple domains.

One of the enduring themes in this historical view of AIED is the emphasis on *personalization*. Personalizing education has been heralded as a transformative equalizer in education (Gardner, 2009), offering a pathway to humanize the learning experience by breaking away from rigid curricula and standardized teaching methods (Midjaas, 1970; Wolfson, 1968).

Rooted in the progressive educational philosophies of John Dewey and Helen Parkhurst (Jenkins, 1998), *personalized learning* emerged as a topic in the literature as early as the 1960s (Trump, 1961). The concept gained traction through its alignment with the individual-focused goals of special education (Adleman, 1971; Blackhurst, 1965; Fischer & Rizzo, 1974) and advancements in educational technology that seemed to make this possibility a reality (Gagné, 1974; Harrison, 1978; Kulik et al., 1976). Today, GenAI is the next step in a long line of educational technologies that have been "guaranteed" to fulfill the promise of individualized learning.

If personalized learning is a *humanizing* endeavor for students (Fischer & Rizzo, 1974; Midjaas, 1970), it is unlikely then that any technology on its own will fulfill this promise. As Vallor (2024, p. 10) observes, GenAI serves as "a mirror of ourselves, not as we ought to be or could be, but as we already are and have long been." Students becoming who they *could be* is a central goal of education generally and individualized learning specifically. STEM education carries additional responsibilities for equipping students to address global challenges, such as climate change, disease prevention and response, and hunger. GenAI by itself cannot fulfill these aims. As such, in this brief, we focus on *individualized teaching* and the ways in which GenAI can support the STEM teacher. Teachers need to be truly kept in the loop, with the GenAI acting as a thought partner or a manager of mundane classroom tasks, thereby allowing teachers to focus their energy and time on connecting with students as whole individuals. This brief explores individualized teaching and how GenAI can enhance teachers' engagement with students.

However, GenAI in its current off-the-shelf form, even when aimed at teachers rather than students, is far from perfect. First, GenAI and LLMs are more broadly considered "black boxes" (Latour, 1987), that is, even the designers, engineers, scientists, and programmers who create these models are not able to fully explain how an LLM provides response Y given prompt X. This opacity and fundamental design as creative, predictive text generation engines lead to LLMs sometimes generating fully unexpected, unhelpful, and false information, a behavior unwisely referred to as a "hallucination" (Berberette et al., 2024). Second, having been trained on an astronomically large corpus of data from the Internet, these models have succeeded in automating biases in our society and depicting the dominant worldview and culture represented by the Internet. Most LLMs are also trained on material on the Internet, copyrighted or not, without permission.

GenAI adversely impacts the health and well-being of the planet, society, and individuals. While some progress has been made in LLM design and green energy initiatives, training and deploying GenAI applications still require enormous amounts of energy and water for cooling, and their carbon footprint is only expected to grow (Crawford, 2024). At the same time, while the “human-in-the-loop” approach—discussed in this research brief and promoted in policy documents, such as those from the U.S. Department of Education, Office of Educational Technology (2024)—is intended to enhance human agency, input, and control in the development and training of LLMs, these practices have a high likelihood of exploitation at all levels. Most, if not all, LLMs are trained and tuned through a human-in-the-loop process known as “reinforcement learning from human feedback” (RLHF). Most of this human feedback is provided by exploited workers in the Global South, particularly Kenya and India, who make less than $2 per hour and are unremittingly exposed to racist, violent, and sexual content, requiring trauma care that is inadequately provided by the companies employing them (Perrigo, 2023; Stahl et al., 2024). It is possible to find LLMs that have been trained only on open access and appropriately licensed materials, reducing ethical concerns about using content without permission, but it is much more difficult to find an LLM that has not relied on exploitative RLHF practices. Even in smaller scale research projects, where it is typical for teachers, community members, and students to engage in RLHF, it is essential to treat them more as *partners* than *participants*, while ensuring fair and equitable compensation for their time and expertise (Ballance & Ripley, 2023; McKee, 2024; Rabinowitz et al., 2024). Extra attention should be paid to the overall well-being of those engaging in RLHF, as the repetitive and tedious nature of RLHF can take a toll on individuals’ mental, emotional, and even physical health. Maintaining clear, consistent, and respectful communication while sustaining their dignity throughout the process provide crucial safeguards against exploitation (McKee, 2024).

There are potentially great benefits to be drawn from GenAI for STEM teaching, but its ethical, environmental, and labor-related concerns should be carefully considered and addressed when embarking on a GenAI research initiative. The National Science Foundation’s (NSF’s) call for RAPID proposals highlights the promise of general-purpose AI technologies and the importance for developing GenAI-based tools and environments that advance equitable learning and inclusive teaching, as well as integrating AIED in ethical, responsible, and effective ways (National Science Foundation, 2023). Additionally, the NSF’s EducateAI initiative underscores the need for research to build an empirical base for AI-powered educational interventions in the age of GenAI.

This research brief reviews early GenAI research efforts and provides guidance on advancing STEM teacher education research, addressing teacher and school reactions, and considering the potential of integrating LLMs to effectively support teachers. This is not a comprehensive literature review but a focused examination of ongoing efforts and the impact and future directions of generative AIED. We begin by providing a brief grounding in pre-GenAI efforts that bear the potential to influence future efforts. We then share the current state of research and practice with respect to AI’s impacts on STEM teaching, focusing on key dimensions of STEM teaching including justice, equity, and inclusion; the teaching process (i.e., curricular decision-making and lesson planning, pedagogy and instruction, and assessing student work and progress); and teacher learning and leadership. We conclude this brief by outlining emerging promising trajectories for research and practice around GenAI and STEM teaching, while also highlighting areas of serious concern and caution.

The Pre-Generative Roots of AI in STEM Teaching

AIED, powered by data and learning analytics, has long promised to enhance accessibility, scalability, effectiveness, and personalization for students through diverse teacher- and student-facing tools. Here, we share a brief account of the history of AIED efforts over the last three decades. Papers such as MacFadyen (2022) and Holmes & Tuomi (2022) share more detailed, systematic reviews of the current state of learning analytics and AIED.

Beginning with the extensive work on ITSs (Anderson et al., 1995; Koedinger et al., 1997) that were built using pre-ML, symbolic AI techniques, researchers have made considerable progress in big data analytics-powered AIED efforts. These approaches involve the use of computational analytics and techniques that use log data from digital environments to provide insights that aid the teaching and learning process (Grover & Korhonen, 2017). At the K–12 levels, AI tools have helped in assessing student learning in formative settings more efficiently through auto-grading student responses and providing feedback. In recent years, these capabilities have extended to assessing more complex responses in K–12 STEM settings (Zhai et al., 2020). Research suggests that well-designed AIED systems can lead to improved student outcomes (Williamson & Kizilcec, 2022) by providing student-facing and teacher-facing support. Student-facing systems have offered personalization to support individualized student outcomes by offering learning support, help (Aleven et al., 2016), and automated feedback (Maier & Klotz, 2022) to students. Such systems also include an analysis of student behaviors, such as engagement or going off-task (Baker & Rossi, 2013), as well as innovations that have the potential to make student learning more engaging and interactive. Studies in K–12 science classrooms have shown how natural language processing (NLP) and data mining techniques can aid in understanding student collaboration (Emara et al., 2021), and how ML–enabled automated feedback can help support students' revision of scientific arguments based on data drawn from simulations (Lee et al., 2021).

Over the years, however, there has been significantly less attention given to developing teacher-facing AIED systems compared to research involving student-facing AI-powered investigations and tools. Tools to support and involve the teacher have typically involved dashboards that serve to provide teachers with insights that allow them to track individual student progress. These insights help teachers to better identify student struggles and, accordingly, better support student learning (e.g., Hutchins & Biswas, 2023b). Some of the most compelling AIED in STEM exemplars that are emerging are systems that are both student- and teacher-facing, such as InqITS (Gobert et al., 2023) and Assistments (Feng et al., 2023). The latter, in particular, is an example of a system that began as a more classic student-centered ITS, but expanded and improved with teacher-facing dashboards and features.

These prior ITSs and AI in education work, in conjunction with emerging projects, provide many lessons for the next generation of AIED efforts involving GenAI. Among the most important lessons learned is that for AIED to succeed in supporting our goals for K–12 STEM teaching, it needs to be human-centered. More specifically, successful AIED systems and tools require thoughtful design and execution, a focus on the processes of student learning rather than just the product (or answer), the provision of contextual feedback to teachers and students, and attention to augmenting rather than replacing human teaching (Luckin, 2025). Additionally, with GenAI we have an opportunity to revisit the idea of personalization in education. This technology allows us to not only improve upon earlier notions of personalizing learning for every student but also to expand it to encompass **personalized teaching** for each teacher.

To understand personalized teaching, we will unpack the positionality of the teacher—and STEM teachers in particular. STEM teachers, as cultural beings, bring a tapestry of culturally and experientially influenced norms, expectations, knowledge, and practices (Garcia Coll et al., 2018; Ikpeze, 2016; Lee, 2012). We define “culture” as Nieto does: “the ever-changing values, traditions, social and political relationships, and worldview created, shared, and transformed by a group of people bound together by a combination of factors that can include a common history, geographic location, language, social class, and religion” (2008, p. 129).

STEM teachers in particular, socially conditioned and academically grounded in the latent assumptions that science, technology, engineering, and math are universal and unbiased, ultimately require that their students engage in a great deal of code switching and assimilation in the classroom—to assimilate into the culture of STEM—in order to be seen as successful (Brown et al., 2016; Kaggwa et al., 2023; Keratithamkul et al., 2020). This success comes at the expense of the students’ authentic selves (Morales et al., 2021). Morales et al. (2021) identify such losses of authenticity to assimilate into the STEM fields as:

> ...counterproductive to the creation of a richly diverse and inclusive scientific community that is prepared to address the questions of our modern world, and more importantly, it is deeply disrespectful and harmful to the BIPOC scientists whom the community boasts about recruiting.

While culture is a group attribute, it is possible to personalize teaching based on culture, promoting inclusion and justice through culturally relevant and sustaining pedagogies (Alim et al., 2020; Ladson-Billings, 1995; Paris, 2012; Paris & Alim, 2017). Just as students have agency to contribute to and transform the learning experience influenced by structural and social conditions (Cook-Sather, 2020; Vaughn, 2020), teachers have agency to react and respond to, shape, facilitate, and transform the learning environment (Ko et al., 2022; Marco-Bujosa et al., 2020; Priestley et al., 2015; Stephenson Reaves et al., 2022) to engage in personalized teaching.

We define “**personalized teaching**” as the **intentional adjustments that teachers, as cultural beings with experiences and backgrounds, make to the learning environment to address the unique needs of each student**. The shifts and adjustments teachers make target the cognitive, social-emotional, and cultural dimensions of each student in the classroom. The goal is to facilitate each student’s growth, development, and learning in a way that is tailored to their individual needs, preferences, experiences, backgrounds, and cultural memberships. This is a key component encoded into curricular and pedagogical frameworks that should be further brought to bear on the design and development of GenAI for teaching, such as Universal Design for Learning (Rose & Meyer, 2000, 2002), the Stanford Neurodiversity Project (Stanford Neurodiversity Project, 2024), Critical Practices for Social Justice Education (Learning for Justice, 2023), and the Culturally Responsive-Sustaining Teaching Framework (NYSED, 2018).

How teachers engage in personalized teaching is influenced by their own experiences, backgrounds, and contexts. Transforming the learning environment through personalized teaching requires the affordances of time, knowledge, skills, and specific dispositions from teachers, as well as a recognition and understanding of the teacher as a cultural being, all of which GenAI has the potential to sharpen, impact, and build upon. Personalized teaching is particularly vital for STEM teachers who seek to engage and empower students, especially marginalized students who face barriers in the learning environment that prevent them from reaching their full potential. These barriers often arise from systemic inequities embedded in structural hierarchies related to racial, ethnic, cultural, and linguistic diversity as well as neurodiversity (Buxton & Allexsaht-Snider, 2017; Ladson-Billings, 2021; Stanford Neurodiversity Project, 2024); disconnects between students’ communities, identities, and lived experiences and traditional curricula (Price & McNeill, 2012; Lee, 2021); and the social-emotional challenges that stem from these misalignments (CASEL, 2024). Personalized teaching requires actively recognizing and addressing these barriers in collaboration with students and their communities.

New Directions in AI for STEM Teaching

To explore the emerging directions for AI and STEM teaching, we follow the major aspects of teachers' work and identify possible new research directions for the intentional inclusion of GenAI. The aspects of teaching we will explore follow:

1. **Justice, equity, and inclusion**. As an integral component of their work, it is essential for teachers to grapple with the question, "How can I use AI to help me facilitate an inclusive STEM learning environment that promotes justice and equity, particularly for marginalized students in my classroom?"

2. **Teaching process:**

   - **Curricular decision-making and lesson planning**. This section addresses the question STEM teachers are faced with, "How can AI help me understand what needs to be taught in STEM and why?"

   - **Pedagogical practices and instruction.** This aspect of teacher work is in service of the question, "How can AI help me in varied pedagogical approaches, activities, and engaging experiences for teaching STEM content and practices?"

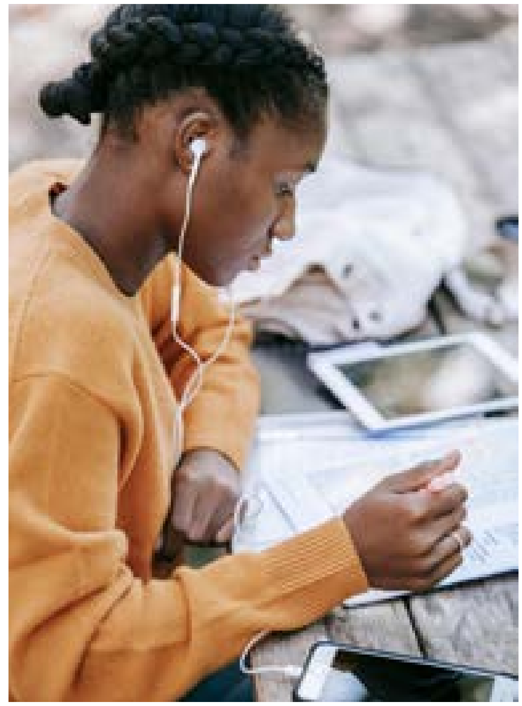
   - **Assessing student work and progress.** This aspect helps teachers explore the question, "How can AI support me in ascertaining where my students are in achieving learning goals (along both cognitive and affective dimensions), what they know, and what they are able to do?"

3. **Teacher learning and leadership.** Typically not considered an aspect of teacher work, but included in expectations by schools and districts of what STEM teachers should do, this aspect addresses the question, "How can AI further teacher growth in STEM teaching and build my capacity for leadership among my peers?"

Questions for Teachers to Consider

- How can I use AI to help me facilitate an inclusive STEM learning environment that promotes justice and equity, particularly for marginalized students in my classroom?
- How can AI help me understand what needs to be taught in STEM and why?
- How can AI help me in varied pedagogical approaches, activities, and engaging experiences for teaching STEM content and practices?
- How can AI support me in ascertaining where my students are in achieving learning goals (along both cognitive and affective dimensions), what they know, and what they are able to do?
- How can AI further teacher growth in STEM teaching and build my capacity for leadership among my peers?

To consider these aspects as a learning environment or system that engages student learning, we must keep in mind that while teachers develop critical *social* relationships with students, their interactions with students in the service of *teaching and learning* are mediated by the curriculum, pedagogical practices, and assessments (Christie & Lingard, 2020; Hayes, 2003; Hennessy et al., 2005; Priestley et al., 2015). Students learn from the teacher *through* those experiences fostered and facilitated by the teacher. At the same time, teachers are themselves students, so they grow and develop professionally through classroom experiences and formal opportunities for training. Ideally, this system is situated on a foundation of justice, equity, and inclusion.



GenAI has the potential to impact each of these three aspects of teaching. GenAI's impact on student learning is explored in depth in *The Potential of Using AI to Improve Student Learning in STEM: Now and in the Future.* GenAI impacts the ***interaction*** between the teacher and students *mediated* by aspects of the teaching process. This impact can be seen in two broad ways: as an **efficiency booster** and as a **thought partner**. As an efficiency booster, GenAI can streamline workflows and increase productivity, creating more time for teachers to engage in personalized teaching. As a thought partner, GenAI can provide feedback on decision-making, help clarify ideas, and suggest new approaches, thereby supporting teachers in developing the knowledge, skills, and dispositions required for effective personalized teaching. We will touch upon these dual roles of GenAI as we discuss each aspect of teaching in detail.

Justice, Equity, and Inclusion

As diversity within schools and the roles of schools and educational systems in upholding oppressive structures, practices, and beliefs are increasingly recognized and politicized, it is a responsibility of all STEM teachers to interrogate their own work to foster more inclusive and more just learning environments by leaning into the existing literature that points to opportunities to engage in self-reflection and self-awareness and to identify issues of bias.

Situating GenAI in Justice, Equity, and Inclusion Efforts

The rapid pace of change in and uncertainty around GenAI opens the door for researchers to investigate, define, and provide pathways through this GenAI moment at the outset to ensure that justice, equity, inclusion, and community cultural capital and wealth are encoded within the orientations and structures of a GenAI system that will assist—rather than replace (De Cremer & Kasparov, 2021)—teachers in planning STEM activities. This rapid and critical response will help to define the educational GenAI research and practice agenda and orient it in a way that promotes a local and context-bound experience, helping to foster cultural sustainability (Alim et al., 2020; Laster et al., 2020; Paris, 2012).

GenAI offers a unique opportunity to address sensitive topics that may be uncomfortable for some teachers—such as implicit biases and deficit-based perspectives and practices. Research suggests that people often react less defensively when interacting with AI agents compared to human counterparts (Candrian & Scherer, 2023). For instance, a teacher seeking to personalize their teaching by supporting and sustaining students' cultural identities through STEM activities, or reflecting on implicit biases and deficit-oriented approaches in their STEM curriculum, might be more open to feedback from GenAI than from a colleague, coach, or supervisor. GenAI can foster a sense of being heard and is adept at detecting and responding to users' emotions (Yin et al., 2024). Additionally, with its core design principle of being helpful to humans (Anthropic, 2024), a GenAI chatbot has the potential to guide teachers in identifying and addressing biases. It can help them adopt an assets-based, culturally sustaining teaching approach in a nonjudgmental environment, allowing them to process and learn without fear of disappointing or offending others. This creates space for teachers to recognize themselves as students, just like their students.

Our commitment to personalized STEM teaching emphasizes the need for GenAI to generate responses that are specific to community cultural wealth and capital (Solorzano & Yosso, 2001; Yosso, 2005) and to the unique assets and needs of individual students (Cioè-Peña, 2022; Rose & Meyer, 2000). This approach contrasts with generic "unbiased" or "color-blind" outputs, which often fail to address systemic inequities (Slota et al., 2021; Watkins, 2021). GenAI relies on vast datasets to produce statistically optimized results and therefore tends to default to "normal" identities—white, monolingual English speakers, upper-middle class, suburban/metropolitan, neurotypical, able-bodied, cisgender, and straight. Such biases make it poorly equipped to handle "edge cases" that represent identities and communities outside of these norms or to address the complexities of intersectional identities (Johnson, 2024). Emerging evidence, however, suggests that GenAI can support teachers in developing ethical, culturally responsive, and context-specific approaches to teaching and learning (Alasadi & Baiz, 2023; Colace et al., 2020). This potential depends on training GenAI with an appropriate corpus and fine-tuning it effectively. With intentional prompting and scripting, GenAI can draw on training data that amplifies the voices, practices, experiences, and worldviews of multicultural, Indigenous, and other historically marginalized communities.

Examples of GenAI for Justice, Equity, and Inclusion Efforts

In their NSF-funded RAPID effort, Price et al. (2023) are engaged in constructing, training, and fine-tuning a GenAI LLM chatbot to assist teachers in developing culturally responsive (Brown et al., 2022; Gay, 2018) and culturally sustaining (Alim et al., 2020; Paris, 2012; Paris & Alim, 2017) STEM curricula and activities. In addition to utilizing national and state STEM standards, summaries of scholarship on culturally responsive and sustaining curricula, and open source libraries of STEM activities as training materials, the LLM will be trained on the voices of intersectionally diverse families, community members, and educators. Drawing on a values-centered design research methodology (Calo et al., 2021; Hendry et al., 2021), the participants are prompted through a multi-step process to relate stories of Gen AI supporting teachers in enacting just and inclusive practices. All participants centered the importance of honor and love toward students beyond the cognitive and disciplinary growth of students, prompting additional training materials that focus on the ethic of caring and respect (Fritzgerald, 2020; Noddings, 1988, 2010). In promising initial results, the pilot version of the AI chatbot is providing general guidance in response to prompts that promotes a culturally aware approach to STEM teaching, prompting the teacher to consider the social-emotional experience of students. This AI chatbot provides this guidance while citing research-based sources that support this approach.[1]

This pathway for supporting diversity, equity, and inclusion efforts in STEM teaching through the use of GenAI requires an intentional effort to identify appropriately focused training data, tuning protocols, and prompt engineering frameworks.

Research Trajectories for GenAI in Justice, Equity, and Inclusion Efforts

In addition to the particular focus on culturally responsive and sustaining approaches to STEM teaching, the participants in Price and colleagues' (2023) project also had the opportunity to define the roles and responsibilities of a chatbot in the classroom in innovative and surprising ways. Researchers and learning scientists can only go so far in anticipating appropriate approaches to support teachers, students, and communities. It is increasingly clear that interdisciplinary teams, as well as members of the communities that the research is intended to support, should be at the table throughout the entire research process cycle (Watkins, 2021).

This focus on partnerships between researchers, teachers, and communities to support STEM teaching that addresses the generational legacies of racism, classism, and ableism (to name a few) is critical for supporting effective and meaningful STEM teaching. GenAI has potential to support these efforts, as explored in the *Toward Ethical and Just AI in Education Research* brief (Barnes et al., 2024). Because of the centrality and critical need for just, equitable, and inclusive STEM teaching, this section is first to set the stage and inform the remaining sections of this research brief.

Curricular Decision-Making and Lesson Planning

STEM teachers make thousands of decisions and judgments each school day. The decisions and judgments they make around curriculum and lesson planning, however, have direct bearing on the learning in which students engage, setting them up for success or failure. STEM curricula and lessons not only provide a set of activities that students do, but also highlight the depth and breadth of STEM content and practices that students learn and

---

[1] While GenAI chatbots have a tendency to "hallucinate" and create research studies through a probabilistic algorithm to support their responses, this AI application has not; it has only drawn on actual published research reports. It has, however, stretched these research reports beyond their scope and set of evidence-based claims. Innovations in Retrieval-Augmented Generation (RAG) have demonstrated a decrease in hallucinations and an increase in the ability for the GenAI to explain where information is coming from (Agrawal, et al., 20204; Alcaraz, 2024; Couldwell, 2024).

embody the values through which STEM teaching and learning are enacted (Hu & Guo, 2021; Ko et al., 2022). GenAI has the potential to support STEM teachers in curating, designing, and adapting STEM lesson plans and activities (Karpouzis et al., 2024; Wang, 2023). In addition, as GenAI can accomplish some of the mundane tasks required for classroom management, the teacher can engage GenAI in designing supportive learning environments through uniquely tailored activities and curricular pathways.

There is a long tradition of research into teachers' curricular decision-making and lesson planning (Ball & Cohen, 1996; Ben-Peretz, 1975; Borko et al., 1990; Remillard, 2005; Schwab, 1973), particularly in STEM subjects. Much of this research focuses on (1) how teachers can support *student learning* through curriculum design (Anderson & Tully, 2024; Cherbow & McNeill, 2022), (2) *teacher learning* by developing knowledge and competencies through the planning process (Amador et al., 2022; Lertdechapat & Faikhamta, 2021; McFadden & Roehrig, 2017), and (3) the *disruption of the typical hierarchical approach* in STEM curriculum design (Chiu et al., 2021; Stephenson Reaves et al., 2022). Curriculum design and lesson planning are frequently part of the work of STEM teachers that neither students nor the general public typically see, requiring additional time and energy during and outside of the regular school day.

Examples of GenAI for Curricular Decision-Making and Lesson Planning

GenAI provides opportunities to support teachers in all of these areas of curriculum design and lesson planning. Through their NSF-funded RAPID AI project, Wang (2023) is currently exploring the ways in which GenAI can support teachers in creating STEM curricula based on the WGBH Educational Foundation's STEM learning resources. The project will provide teachers with "LLM building blocks," allowing teachers to guide the GenAI in a way that is consistent with their goals for the lesson. Karpouzis et al. (2024) are designing an NLP "digital assistant" for educators to provide assistance in tailoring lesson plans according to factors such as the teacher's goals and student characteristics. It should be noted that GenAIs lean toward expressing systemic and latent biases (Zhou et al., 2024) and tend to provide generic—rather than specific—responses to complex problems and issues in areas as disparate as the digital humanities and aviation accident investigation (Guo, 2024; Ziakkas & Pechlivanis, 2023). The STEM curriculum and the Next Generation Science Standards are rooted in the complex interweaving of specific conceptual ideas, crosscutting concepts, and practices. The tendency toward the generic can lead to cursory exploration of content and lax approaches to practices. Mathematics education requires a similar focus on details; Sakamoto et al. (2024) found this generic tendency to be problematic for teachers in a cross-cultural evaluation of an AI-generated math lesson plan. Any effort to use GenAI for the creation of curriculum should be examined closely and treated as a broad strokes outline, with human teachers guaranteeing the depth and rigor necessary for the creation of good STEM curricula.

There are structural issues in curriculum design and lesson planning that GenAI cannot fix, such as societal biases and the uncompensated and extensive demands on teachers' time outside the classroom when much of this planning takes place. Generative AI, however, can impact the efficiency of teachers in many tasks in the curriculum design and lesson planning process (Karpouzis et al., 2024), supported by research in other fields (e.g., Al Naqbi et al., 2024; Brynjolfsson et al., 2023; Noy & Zhang, 2023).

Research Trajectories for GenAI for Curricular Decision-Making and Lesson Planning

As noted above, there are trade-offs to this boost in efficiency. A promising direction in the research is viewing human-AI interactions as ongoing dialogues and co-constructive processes so that teachers and educators provide feedback to the GenAI to ensure that it is providing effective and relevant—not just efficient—responses to teachers' queries (Eriksen & Eriksen, 2023; Muller et al., n.d.; Sun et al., 2024; Treleaven & Brown, 2024). In an ideal world, time saved in lesson planning would allow teachers to devote more time to attending to the more human aspects of relationship-building and addressing individual student needs. That said, providing teachers more time to focus on constructing the resources, activities, and lesson plans necessary to engage the individual students in their classroom may be more sustainably addressed through broad changes in policy and societal norms to reorganize the expectations placed on teachers. A promising trajectory in research on STEM curriculum development and lesson planning, however, is exploring the ways in which teachers and GenAI learn from each other. Such as exploration may not only improve efficiency in the long run but may also result in context- and culturally relevant and sustaining STEM curricula to better support the individual students in their classrooms.

Pedagogical Practices and Instruction

Much of the school day for STEM teachers is spent interacting and engaging with students in multiple ways, from whole class direct instruction to group work to one-on-one coaching and tutoring. This section highlights research that has illuminated the ways in which AI can support STEM teachers in their pedagogical and instructional practices, the management of their classroom, and opportunities for new areas of research. As noted throughout this research brief, the emphasis in this section is on *assisting* the teacher rather than *replacing* the teacher.

There are a number of different ways in which GenAI can support teachers in the active pedagogical moves in the classroom. First, GenAI can support teachers by providing one-on-one or small group individualized coaching to students, building on past work around tutoring systems. The second is through supporting teachers in identifying and utilizing appropriate instructional or pedagogical moves in the classroom based on interactions with students. A third area is one that is highly practical, assisting the teacher in classroom management, potentially allowing the teacher to focus on engaging in personalized teaching methods to support the individual students in their classrooms. We share some examples below.

Examples of GenAI in Support of Pedagogical Practices and Instruction

Through a self-study process of examining interactions with ChatGPT, Cooper (2023) identified several ways in which GenAI can help direct teachers through their instructional decision-making. However retrospectively or prospectively, Cooper found that GenAI was able to engage in conversations around redirecting the teacher toward more student-centered pedagogies, reminding the teacher of their duty to care for all students and to develop tools for specific purposes, such as activities and rubrics, while teaching. Chen (2023) noted the ability to use AI image generators to create images that would provide a hook into a lesson or topic that may not be available otherwise. However, caution and vigilance are necessary. Edouard (2024) found AI image generators to embody and perpetuate racial and cultural biases, and thus hinder culturally relevant designs when Black students used them as a co-creative tool for designing game characters. Demensky & Liu (2023) presented an NLP-based automated tool to provide instructors with feedback on dialogic teaching practices and teacher responsiveness, including student engagement, talk time, and questioning strategies in an online introductory programming course for high school students.

Research Trajectories for GenAI in Support of Pedagogical Practices and Instruction

Beyond the possibilities, it is necessary to think through the infrastructures that may facilitate these processes. A dashboard-based system backed by GenAI is a promising model for the field (Hutchins & Biswas, 2023a, 2024). By integrating co-designed learning analytics, data visualizations, teacher reflections and observations, and AI supports, GenAI may be able to provide a robust platform of just-in-time support for teachers to reflect on in the moment and consider next instructional steps. It will be critical to have teachers in the loop to help design and fine-tune the system so that it can reliably provide this type of support.

Classroom management was a surprising—although significant—finding for the use of AI in STEM teaching. Through a community-engaged and values sensitive design process based on the works of Bang & Vossoughi (2016) and Hendry et al. (2021), Price et al. (2023) facilitated listening sessions with teachers and families at an elementary school in a large city school district. Participants shared stories of AI serving, in many ways, as an instructional assistant. While the classroom teacher worked with individuals or groups in a STEM classroom, an AI engaged the other students in the class in topical conversations and activities. This is a tacit recognition that teacher-student engagements are most effective as one-on-one discussions or in small groups. This GenAI provided whole class coverage while the teacher engaged in focused teaching. Participants also identified another potential role for GenAI: assisting teachers in checking in with students on their social-emotional well-being. This is particularly significant as research consistently shows that students' experiences outside of school profoundly influence their engagement and learning in the classroom (Aronson & Laughter, 2016; Blanchett et al., 2009; Nespor, 1997; Noddings, 1988). These imagined futures, while potentially controversial and not without unforeseen consequences, envision classrooms where a single human teacher is supported by multiple GenAI agents. These agents could address various needs of both students and the teacher, supporting personalized teaching and helping maintain a positive, effective learning environment.

What was made crystal clear in the stories that emerged from participants in the Price et al. study (2023) was that teachers were eager to do the job they were hired to do: teach.

GenAI has the potential to assist teachers in novel and innovative ways through the practicalities and practices of teaching in the dynamic space that is a classroom. Ideally, this assistance would yield greater efficiency in the persistent tasks that teachers are required to accomplish. Indeed, there is evidence that teachers are utilizing GenAI for these very reasons (K. Johnson, 2023, 2024). There is plenty of evidence, however, that technologies intended to create efficiencies and therefore provide the user with more time for other pursuits have led to unintended consequences or backfired, a phenomenon known as the *productivity paradox* (e.g., Aggarwal-Schifellite, 2017; Barros, 2023; Hajli et al., 2015). While research on the impacts of GenAI on teacher efficiency is important, it should be done within the context of societal expectations and existing accountability structures.

Formative and Summative Assessments

Classroom-based formative and summative assessments are crucial for providing important feedback to teachers that illuminate where adjustments need to be made to curriculum and teaching strategies. They also provide feedback to students about where they are in the learning journey toward stated learning goals.

Situating GenAI in Formative and Summative Assessments

As with AI-related research in the previous decade, a compelling use case for GenAI is to aid in automated scoring and feedback for students and teachers. As it happens, many of the recent "pre-Generative AI" AI/ML in education efforts, especially in science education, have been focused on assessing student learning and specifically for automated scoring of open-ended responses by students as part of scientific argumentation and explanation (Zhai et al., 2020). Other efforts have focused on studying learning processes through examining trace data from digital learning environments as a type of formative assessment (Gobert et al., 2013; Grover et al., 2017; Shute & Ventura, 2015). This section provides insight into research on STEM assessment using GenAI tools as well as potential promising research trajectories.

Examples of GenAI in Formative and Summative Assessments

Multiple science education research efforts are leveraging these new technologies to aid teachers in the analysis of long-form, open-ended responses of students' scientific argumentation and explanation. Such explanations are time consuming to assess manually, and they either do not get comprehensively assessed by teachers or are not processed in a timely manner to provide the intended timely formative feedback to teachers and students. Such assessments also present challenges as they involve evaluating students' ability to apply multiple dimensions of scientific knowledge to comprehend and analyze phenomena (National Research Council, 2011; Zhai et al., 2020). LLMs such as Google's BERT[2] (Devlin, et al., 2019) were precursors to today's LLMs (e.g., ChatGPT, Gemini, and Bard) and were able to provide support for and accomplish these tasks. These models excel in this area since the emphasis is on the analysis rather than the generation of text, which is the domain of GenAI.

[2] BERT stands for "Bidirectional Encoder Representations from Transformers."

In other research, students' proficiency in scientific explanation was categorized by combining Google's BERT with an ontological framework tailored to a contextualized science assessment. Their findings revealed that although pretrained language models, such as BERT, improve performance on language tasks in education, using ontology-based systems to identify domain-specific terms and replace them with related sibling terms in sentences can notably enhance the accuracy of classification models. In a similar vein, Cohn et al. (2024) use human-in-the-loop NLP to automatically score short answer questions in formative "check-ins" in a curriculum involving the integration of science, computing, and engineering concepts. Specifically, they employ a chain-of-thought (CoT) prompting technique, which encourages step-by-step reasoning in the model's responses. This is combined with active learning on the part of the model, where it is improved by incorporating input from teachers and students. This process refines the model's scoring accuracy. Another promising aspect about this research is that these LLMs can score formative assessments regarding scientific processes as expressed in diagrams and images (with arrows) by citing evidence from students' responses and tying it back to the rubric to help guide scoring decisions. This extension to traditional CoT prompting of using the reasoning chains to cite evidence from students' responses, both improves the LLMs' scoring abilities and allows LLMs to provide feedback to teachers and students explaining why (or why not) students are awarded points based on the evidence from their responses and the rubric. Both Wang et al. (2024) and Cohn et al. (2024) report on high accuracy rates when compared with human scoring.

In addition, researchers are engaging in research on GenAI technologies to help teachers not only with achieving greater efficiency and accuracy, but also in accomplishing things teachers simply could not do before. In an NSF-funded RAPID project led by Grover & Clarkson (2023), high school mathematics teachers from rural, suburban, and urban Indiana use ALICE, an LLM trained on the open source WeBWorK (Gage et al., 2001) library of math problems. ALICE generates technology-enhanced assessments (TEAs) that are auto-gradable, interactive, and randomized (isomorphic), features that support classroom formative assessment to aid the teaching and learning process. ALICE takes prompts in natural language to generate the WeBWorK code for the corresponding TEA along with hints and a teacher solution. Such code is normally written by programmers, effectively excluding K–12 teachers from the process of creating WeBWorK TEAs. Preliminary findings show that teachers using ALICE feel empowered and excited about creating seemingly endless variations of problems, and they are leveraging ALICE's "creativity" to create problems that are relevant to their students' contexts (Grover et al., 2024). However, the more interesting findings from teacher interviews pointed to a whole host of factors that influence a teacher's experience working with an LLM for math problem generation—usefulness in teaching, prompt creation, and refinement process; student use and reactions; viewing AI as a thought partner; comparison with their usual approach (to assessments); and attitude toward GenAI generally and its future use. These findings suggest that the question is not whether a teacher will use GenAI or not, but the question should be how teachers will use Gen AI in context. The project highlights the importance of understanding the science of teacher-AI teaming and domain-specific prompt engineering.

Research Trajectories for GenAI in Formative and Summative Assessments

An emergent theme appears to be the fine-tuning of GenAI responses through techniques such as CoT and few-shot learning (an approach to training GenAI through limited, rather than broad-ranging and repeated, exposure to training data) with the goal of guiding the LLM to provide more targeted support in assessing student learning. Another emergent theme is the design and training of domain-specific GenAIs or otherwise augmented GenAIs. This approach has the potential for supporting diverse students and the implementation of culturally relevant pedagogies in the classroom. Given that the current corpus of general-purpose GenAI LLMs (such as those from OpenAI, Google, and Meta) have been trained on the Internet as their dataset, it is fair to assume that these LLMs do not represent Indigenous peoples' and other cultures' knowledge and ways of knowing and doing STEM due to the statistical centering and outlier reductions that take place. We can take a leaf out of the ongoing approaches in the research projects described above to design and train custom LLMs or fine-tune or otherwise augment training data for GenAIs to fill in the gaps and make them work for all our students and in diverse settings. Such GenAIs can aid in personalizing teaching and learning by automating the design of assessments that are keyed toward individual students. As an example, GenAI can ensure students are not all answering the same multiple choice questions, but rather questions that are geared towards them individually. These GenAI systems also provide options to allow assessment to be more multifaceted, multimodal, and multidimensional.

In-Service and Preservice Teacher Learning and Leadership

STEM teaching requires continuous learning to stay abreast of new discoveries in the STEM fields and innovative pedagogical and curricular practices to address a dynamic classroom and the needs and dispositions of individual students. It is also a profession that requires maintaining licensure and sustaining a rigorous induction process. These requirements place paramount importance on professional development (PD) and learning. In addition, it is often expected that STEM teachers will rise to fulfill peer leadership roles, particularly since there are institutionally few opportunities for advancement within the role of the STEM teacher. This section explores the current state of research and the possibilities for AI in helping in-service and preservice STEM teachers grow and develop capacity for leadership.

Situating GenAI in Teacher Learning and Leadership

Up until this point, the role of GenAI has been focused primarily on the teacher's role in changing and transforming the learning environment. However, the focus on teacher learning in this section now considers the educator as a whole human being with a sense of invested agency in the learning environment. GenAI, then, has the potential to be incorporated into learning environments where the teachers are beneficiaries of personalized teaching, which then serves as a foundation for their professional growth. Research has shown that while sustained professional learning that explores, uncovers, and stretches teachers' experiences in concrete situations are effective in teacher learning and growth (Huang et al., 2022; Miller et al., 2021; Sims & Fletcher-Wood, 2021), many in-service teacher learning opportunities fall short of meeting the needs and challenges of in-service teachers (Dede et al., 2016; Gould, 2008; Miller et al., 2021). There are also access gaps to quality PD for teachers at under-resourced and rural school districts (Copur-Gencturk et al., 2024).

Examples of GenAI in Teacher Learning and Leadership

Chiu (2023) identifies several areas that GenAI can impact the outcomes and content of both in-service and preservice teacher learning opportunities, including 1) supporting the development of teacher facilitation and leadership skills, 2) promoting interdisciplinarity, and 3) fostering teachers' AI literacy. In a randomized controlled study, Copur-Gencturk et al. (2024) found that GenAI and LLMs serving as virtual facilitators providing real-time feedback to teachers are significantly more effective at improving student outcomes than the traditional models of in-service teacher PD. In essence, preparation for *personalized teaching* with GenAI can be engendered through *personalized teacher PD*. Such personalization can take into account teachers' backgrounds, their existing teaching philosophies, and their attitudes toward AI (Grover et al., 2024).

This sustained learning may aid STEM teachers in becoming teacher leaders, coaches, and mentors, critical stepping stones for growth in a profession that has few opportunities for advancement. In addition, understanding that there may be unintended consequences and that larger barriers are at play, the efficiencies afforded by GenAI may free up time for teachers to engage in mentorship and collaboration. Also, just as with most new technologies introduced to the classroom, early adopters of GenAI are helping colleagues use chatbots for professional purposes, serving as either informal or formal mentors (Johnson, 2024).

Research Trajectories for GenAI in Teacher Learning and Leadership

To foster changes in preservice and in-service learning opportunities for STEM teachers with GenAI, Langran et al. (2024) identified a constellation of factors that require negotiation and navigation for successful at-scale outcomes. These include engaging in ongoing and organic discourse, building momentum for the use of GenAI, and building on existing structures, such as university and state-mandated professional requirements, partnerships with local schools, and expertise from professional associations and experts. These possibilities point to the role of infrastructuring-based research (Hopkins et al., 2013; Hopkins & Woulfin, 2015; Penuel, 2019) in the area of GenAI and teacher learning. While a focus on co-constructing and analyzing infrastructures is important in all areas of this report, it is particularly salient in the area of teacher learning and leadership due to the complex nature of the in-service and preservice teacher learning enterprise, which is shaped by local, state, and national sociopolitical climates and policy initiatives; demographic and socio-ecological contexts; and access to technologies, expertise, opportunities, and offerings. As an organization that provides national standards, PD, and existing teacher learning networks, ISTE can serve as a practical national-scale resource and partner for teachers and educational researchers.

The Way Forward for AI in STEM Teaching: A Complex Interplay of Opportunities and Risks

In this section, we share additional overarching themes that have emerged from our commentary and current research presented in the previous sections. Embedded in these themes are the following: (1) broader takeaways than those shared in the previous sections; (2) opportunities for future empirical inquiry; (3) opportunities for immediate and future research; (4) guidance that can ensure mindful, just, and productive use of AI in STEM classrooms; and (5) concerns that must be dealt with and addressed if we hope to succeed in this endeavor. In the end, a key goal of this brief is to ensure that future research on AI in STEM teaching is purposeful and mindful of the issues described here.

Opportunities: Promising Research Trajectories

GenAI as an Artificial Teacher's Helper

For the last three or more decades, teachers have borne the brunt of having to incorporate into their classrooms the endless flow of purportedly promising "edtech"—computing-enabled technologies that, more often than not, have been developed by technologists without meaningful collaboration with teachers or engagement in classrooms. The hype cycles accompanying this unending array of new technologies tout how the "next new thing" will be the silver bullet to address the challenges of teaching STEM subjects to all students. In the typical hype cycle, pressures from administrators force teachers to invest time and effort to learn about and leverage these shiny new tools. After decades of witnessing the familiar cycle where the promises of educational technology often under-deliver or even cause setbacks, it is time for policymakers, administrators, technology developers, and researchers to acknowledge the immeasurable importance of teachers in the classroom, as consistently highlighted in this research brief. The necessary and invaluable human relationships they bring to the learning equation cannot be overstated. AI tools must support, not supplant the teacher. We cannot minimize the importance of the human connection, and neither can nor should we put our classes on auto-pilot. Classrooms should not be conceived of as potential autonomous driving vehicles. Even in settings where students are being provided individualized guidance from AI, the teacher must always have their hands on the wheel.

We do not believe that GenAI necessarily makes better decisions, but it can be part of a toolkit that allows teachers to make better decisions. There is the need to recognize the monumental task of teaching and the many challenges STEM teachers face in their day-to-day work. There is also a need to develop GenAI tools that are teacher-facing and are consciously designed to help teachers alleviate the challenges they face in enacting good STEM teaching. There are plenty of gaps that can be viewed as opportunities to create supportive tools to aid teachers in their STEM-teaching activities and classroom practice. Such tools must be co-developed and co-designed with teachers (as described in more detail below), in addition to balancing innovation and responsibility as described in the U.S. Department of Education's Office of Educational Technology's (2024) guidelines for developers.

With the mantra of "always center educators (ACE) in AI" as a foundational guiding principle (U.S. Department of Education, Office of Educational Technology, 2023), we share additional themes and concerns for AI in STEM teaching that present opportunities for the NSF Discovery Research PreK-12 (DRK–12) and other education research communities to address.

Contextualizing and Illuminating the "Black Box" of AI in STEM Teaching

In partnership with computer and data scientists, educational researchers have an opportunity to address the shortcomings of GenAI, which have real-world and material impacts on its use by STEM teachers. For such systems to be fully effective, the black box of GenAI must be fully open to examination throughout the design and development process.

Even though GenAI typically responds to teaching situations in a very generic manner, as discussed earlier in the Curricular Decision-Making and Lesson Planning section of this brief, it is important to remember that teaching is always contextual. As a recent report on human-AI teaming suggests (National Academies of Sciences, Engineering, and Medicine, 2022), research is needed to determine improved methods for supporting collaboration between teachers and AI systems in shared functions. This includes supporting human teachers working with AI systems at multiple levels of automation and determining methods for maintaining or regaining situational awareness when working with AI systems at high levels of automation (as represented by former 1-1 cognitive tutoring systems).

Currently, most promising emerging strategies for designing socio-technical systems that leverage GenAI LLMs in STEM classrooms appear to involve fine-tuning, augmenting, and/or customizing LLMs to assist teachers in developing STEM activities, assessments, and materials that are not only pertinent to the topics at hand but also culturally relevant and sustaining (Alim et al., 2020; Ladson- Billings, 1995; Paris, 2012; Paris & Alim, 2017). Such socio-technical systems also involve human-AI interactions as ongoing dialogues and co-constructive processes so that teachers provide feedback to the GenAI to ensure it is providing effective and relevant—not just efficient—responses to teachers' and students' queries. A critical, prudent strategy will be to provide teachers with tools and the know-how to perform the last mile of fine-tuning and contextualization of the LLMs they use in their teaching.

In a similar vein, there need to be investigations of how to support teachers in enacting the human-AI-human pathway, where AI use is initiated by teachers, and the output is reviewed, revised, and refined before being used in the classroom. A promising approach in this space incorporates finely tuned language models and/or draws from specialized data sources, such as knowledge graphs (Rathle, 2024), so that the situational knowledge of the teacher, students, and classroom context are considered in GenAI's outputs of lesson planning, assessments, and pedagogical strategies. Whether such customized versions of foundational off-the-shelf LLMs will authentically serve the needs of STEM teaching or whether content- and context-specific small language models (e.g., Liu et al., 2024) will win the day is an area ripe for research. Each of these approaches opens up the development process to teachers and researchers, making the black box more transparent and shining light on the process of preparing and training GenAI systems.

Co-design and Participatory Design with Teachers and Communities

One of the deep challenges with GenAI is ensuring that the LLMs, algorithms, and training data fit the social, cultural, and geographic contexts in which it is employed. Utilizing the typical "more data will make it better" approach tends to contribute to flattened, color-blind, and context-independent responses by AI chatbots (Price, 2024). It is critical then to ensure that teachers and families are involved in the design and development of GenAI-based tools specific to STEM teaching. Community-engaged and critical participatory design research methodologies (e.g., Bang et al., 2016; Bang & Vossoughi, 2016) are promising avenues for ensuring that the voices and perspectives of educators and families are included in the design and development process. Calo et al. (2021) provide such a methodological framework specifically designed to elicit possibilities for AI in context. This methodology guides participants through a structured process with scaffolding to construct a culturally embedded short story demonstrating the activities and impacts of AI in context. The process involves not only cognitive exercises, which include a mix of questions and frameworks to illuminate and relate the stories, but also the incorporation and construction of material artifacts that highlight the importance and durability of their values and ideas.

The inclusion of educator and community voices can also help move the conversation about the design and development of GenAI tools from a deficit-based approach to an asset- or strengths-based approach (Ocumpaugh et al., 2024). Of critical importance is to delve deeper than simply "leveraging assets," as described in most asset-based approaches. Delving deeper into the design and development of GenAI tools for teaching involves efforts to sustain cultures (Alim et al., 2020; Paris, 2012) and pushing back at the normalizing effects of disciplinary STEM learning while providing opportunities for STEM students to bring their cultures and identities with them, carried, supported, and strengthened in the STEM learning process.

A technical solution to this diverse and inclusive approach to constructing GenAIs is by populating vector- and graph-based databases with culturally sustaining knowledge, worldviews, and practices of marginalized communities and cultures, thereby making them available to GenAI systems (IVOW, 2018, 2020, 2021). Providing GenAI applications with this form of structured "data" can increase the likelihood that the responses of the LLM will reflect these racially, ethnically, culturally, linguistically, and neurologically diverse voices and wisdom, which tend to be flattened and minimized in the statistical processes of LLMs. Such an approach can provide new entry points and forge new pathways in STEM teaching and learning.

In a sense, GenAI can serve as a virtual co-teacher or coach to remind teachers of the diversity embodied in their students and provide supports and strategies for teaching and engaging them. GenAI can be used to support these goals, however, only if teachers, families, and communities are part of the design and development process of both GenAI tools and the databases from the beginning of construction and then throughout the entire process.

Leveraging Past Education Research for More Impactful GenAI

The corpus of research that addresses all the aspects of STEM teaching is not only comprehensive but relevant, even as we move to an age of GenAI in STEM classrooms. One could argue that it is absolutely crucial for GenAI LLMs to encode learning theory and findings from research on STEM teaching so that their outputs reflect the best evidence-based knowledge we have on the issues of justice and inclusion, curriculum and instruction design, pedagogy, assessments, and teacher preparation in K–12 STEM.

For example, the importance of learning trajectories and progressions, as they relate to establishing learning goals and standards, driving formative assessment (Harris, et al., 2023), and supporting STEM classroom pedagogy and lesson planning are well recognized. In our literature reviews for preparing this brief, it was not clear whether LLMs draw from research on STEM learning progressions, and even if they do, how well GenAI applications are able to leverage learning trajectories and progressions. By integrating progressions, AI could identify coherent, interconnected pathways that map out how to develop students' knowledge and skills with targeted instructional supports. This approach is promising because the AI would be informed by learning theory based on previous research on student cognition and learning and structured around disciplinary content and practices.

Curriculum and activity decisions are paradigmatic, relying on either latent or explicit (preferred) models of student cognition. Any LLM used by teachers to generate lesson plans and activities must be transparent as to what models of student cognition it draws on. As mentioned in the Formative and Summative Assessments section, there has been a large body of research conducted on ITSs over the last three decades devoted to understanding and encoding models of student cognition into AI models (e.g., Koedinger et al., 2012). The recent LLM-powered GenAI tools do not appear to be leveraging that prior research. A compelling research agenda on AIED should include investigating how we leverage this prior work to augment and enhance general-purpose GenAI LLMs, and how models of student cognition and learning can be embedded in LLMs.

Many state-of-the-art efforts in education assessment are guided by Item Response Theory (IRT; deAyala, 2009), which is based on assumptions that connect student characteristics with specific content pathways. However, new opportunities with GenAI may enable the identification of novel and personalized pathways missed by IRT in only recognizing one learning pathway as the most appropriate one for all students.

GenAI can also draw on contextual factors, such as social-emotional considerations, and leverage a range of existing frameworks and research, ranging from culturally relevant pedagogy (Ladson-Billings, 1995) to the CASEL[3] framework for supporting social-emotional growth (CASEL, 2024), to help teachers weave in an emotionally and culturally supportive pathway. These contextualized pathways could enable educators to engage in focused, personalized teaching that supports whole-person student learning. For example, a GenAI system could analyze data from various sources to identify students who might be experiencing stress or disengagement. The AI could then suggest specific strategies or resources to the teacher, such as mindfulness exercises, personalized feedback, or collaborative learning activities, that address the emotional and academic needs of the student.

There is a broad and deep foundation of research identifying how the teaching and learning process can confront and counter the structural inequalities faced by racially, ethnically, linguistically, neurologically, and economically oppressed and marginalized students, some of which are explored in this research brief. Now, there are renewed opportunities to leverage GenAI's capabilities in all aspects of STEM teaching—through curricular decision-making and lesson planning, pedagogical practices and instruction, and teacher professional learning and leadership. These advancements can help imagine, facilitate, and implement STEM education environments and experiences that eliminate structural barriers to success, providing multiple entry points and pathways through STEM for a more diverse, richer, and deeper STEM field.

Risks: A Critical Appraisal of Generative AI's Promise

Whenever AI is used in the classroom, there must always be heightened concerns for the safety, continued learning, and well-being of the humans using and being supported or guided by the AI tools and systems. As the most immediate and visible, concerns about plagiarism and cheating are often the first raised by educators (Lee et al., 2024; Wignall & Hart, 2024). Such concerns are warranted: cheating impacts learning (Malik et al., 2023) and overall moral development over time (Chance et al., 2011).

However, the concerns should not end there. Issues of ethics and bias in AI also need to be at the forefront of the concerns that educators, researchers, and policymakers grapple with. It is important to move beyond surface-level concerns for ethical uses of and unbiased approaches to GenAI and focus on the promotion of justice and equity. While including matters of student and teacher privacy and data collection, a recent CADRE[4] brief (Barnes et al., 2024) addresses these concerns by articulating a framework for ethical and justice-oriented practices, as well as tools and strategies that can be adopted by education stakeholders as they integrate AI into K–12 classrooms. Researchers and educators should also look to guidance from the *Responsible AI and Tech Justice: A Guide for K–12 Education*[5] and other similar frameworks to ensure that justice and ethics are always front and center.

---

[3] CASEL stands for "Collaborative for Academic, Social, and Emotional Learning."

[4] CADRE stands for Community for Advancing Discovery Research in Education.

[5] See Responsible AI and Tech Justice: A Guide for K–12 Education, published by the Kapor Foundation, https://kaporfoundation.org/wp-content/uploads/2024/01/Responsible-AI-Guide-Kapor-Foundation.pdf

Data Colonialism and Appropriation

While recognizing the potential of GenAI to promote more just and inclusive STEM teaching through the inclusion of marginalized peoples in the design process, it is important to remember that many oppressed and marginalized communities are (rightfully) suspicious of such efforts to collect data from them (Carroll et al., 2022; Klassen & Fiesler, 2022; Scharff et al., 2010). These data have often come at the expense of the health and agency of the communities in which the research has been conducted, providing few opportunities for these communities to have a say in the use, scope, and outcomes of the data they provided. Frequently, these new understandings have been utilized to help *other* communities while ignoring the communities that provided the insights and data. In effect, the data are colonized and appropriated by researchers for their own purposes and benefits (Couldry & Mejias, 2019; Thatcher et al., 2016).

In addition to asking the important question of whose voices GenAI training data represent, it is critical to enter into respectful data-sharing relationships with individuals and communities and to focus as much (or more) on issues of data access, fairness, processes, and sovereignty as on issues of data collection protocols and analytic rigor. The U.S. government's newly released guidelines on designing AI tools for education (U.S. Department of Education, Office of Educational Technology, 2024) provides counsel for AIED developers in such situations (e.g., shared trust, identifying and managing risk, and responsible innovation). Carroll et al. (2022) provide a more in-depth and specific framework for entering into data-based relationships by pairing the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles with the CARE (Collective Benefit, Authority to Control, Responsibility, and Ethics) data principles.

Idiosyncrasies and Misinformation

Other legitimate concerns pertain to the idiosyncrasy of the outputs GenAI weaves together. They are given to inaccuracy and can be used to reinforce and extend misinformation and disinformation campaigns. While researchers and engineers investigate *technological approaches* to dealing with these issues, it is also important to invest in *human infrastructure* and *capital* to mitigate the risks involved in this known very consequential bug of GenAI. One of these investments is developing a GenAI literacy specifically targeted at STEM teaching by incorporating such efforts into preservice teacher coursework as well as PD for in-service teachers. Such an effort would involve building teachers' understanding of how GenAI works and helping teachers recognize not only inaccuracies, but also generic, surface-level content and responses as discussed in the curriculum and pedagogy sections above. This is not a stop-gap effort to be used until a technological fix is employed, but rather a long-term and sustained learning process to ensure teachers—and society—do not become reliant on GenAI while remaining blissfully ignorant about the technology. These investments are a pre- or co-requisite for using GenAI tools in the ways described in this brief.

Teacher Efficiency as a Powerful Idea and Mythical Quality of AI

The promise of improving teacher efficiency in completing routine tasks is a thread that runs throughout this brief and the research on GenAI in STEM teaching. Efficiency is a "promise" of GenAI, one that is greatly desired by the overworked and underappreciated professionals we refer to as STEM teachers. As research that pays attention to the productivity paradox makes clear, however, GenAI's contribution to teacher efficiency is neither direct nor assured. It is critical to recognize these assumptions around boosting efficiency as a "*powerful idea*" (Postman, 1998)—an idea "baked into" the design and marketing of the technology. This leads to the possibility of efficiency taking on a "*mythical quality*" (Postman, 1998) of AI, a characteristic that is integral to the way the technology operates with cursory empirical evidence to back it up. Exploring the possibilities of efficiency as a component of GenAI in STEM teaching is an exciting and potentially fruitful trajectory, and GenAI has the potential to free up teachers' time, thereby providing the opportunity for teachers to engage in more personalized teaching or simply making their lives more balanced and saner.

Research on the promises of efficiency, however, needs to be done within the context of other efforts, such as identifying the "*tradeoffs*" (Postman, 1998) of GenAI's impacts on teacher efficiency. An increase in efficiency doesn't necessarily lead to better, deeper teaching, nor is the authentic fulfillment of teachers' professional selves guaranteed. GenAI's impact is "*ecological*" (Postman, 1998), meaning that the effects and consequences of GenAI are difficult to accurately predict and may lead to outcomes that are unexpected. An overreliance on efficiency also sidesteps the notion that the benefits of GenAI are "*unequally distributed*" (Postman, 1998) across the educational landscape.

Conclusion

It is difficult to predict with any certainty the true nature of the impact GenAI will have on STEM teaching in either the near or far future. As researchers—and optimists—we believe GenAI's most promising role for STEM teaching is as an integrated component of the educational ecosystem, accomplishing mundane and auxiliary classroom tasks and serving as a thought partner for teachers. In this way, GenAI contributes to an educational environment where personalized STEM teaching becomes increasingly likely—but not guaranteed.

To increase this likelihood of potential positive impact, research and development should focus on ***deepening***, ***broadening***, and ***opening*** the possibilities of positive impact by GenAI on STEM teaching. By **centering and including teachers** (and by extension, families and communities) **in efforts to design, develop, and evaluate GenAI**, we ***deepen*** the potential positive impact of GenAI on STEM teaching. GenAI systems are then designed specifically to interact with teachers on issues that are critical to the teaching profession, allowing for deep interaction and engagement.

By **ensuring diverse training corpora and investing in teachers' GenAI literacy practices**, such as teacher and student prompt engineering and critically evaluating GenAI outputs, we ***broaden*** the potential positive impact of GenAI on STEM teaching. Our world, our nation, and particularly our educational system are becoming more diverse, with educators doing their best to meet the needs and experiences of each student. Incorporating training data representing the diversity in the classroom—and at local rather than global scales—will broaden the appeal and utility of GenAI for STEM teaching. Meanwhile, preparing preservice teachers and supporting in-service teachers to engage with GenAI intentionally and critically will allow teachers to use GenAI more effectively to support personalized teaching.

By **specializing, contextualizing, and paradoxically, limiting the roles of GenAI**, we ***open*** the potential positive impact of GenAI on STEM teaching. Major AI companies, such as OpenAI, Anthropic, Google, and Meta, seem focused on creating a single, unified all-purpose GenAI system—a "superhuman" intelligence capable of efficiently tackling any task, regardless of size or complexity (Griscom, 2024). However, this research brief highlights the effectiveness of GenAI when applied to specific tasks within defined domains, such as helping teachers interpret and respond to assessment data or serving as a thought partner in activity development. GenAI systems designed for these specific limited roles can be developed, trained, evaluated, and researched more efficiently and thoughtfully than the quixotic pursuit of an "artificial general educational intelligence" system. This specialization and scope limitation will actually broaden the potential for GenAI to meaningfully impact STEM teaching and foster an educational environment conducive to personalized instruction.

While there is the potential to deepen, broaden, and open potential positive impacts for GenAI on STEM teaching, it is crucial to consider the likelihood of an ***unequal distribution of benefits*** within the educational ecosystem. Historical precedence suggests that schools and STEM classrooms that can benefit the most from GenAI support may, in reality, benefit the least in terms of student outcomes. Teachers in “underperforming schools”—which often serve poor urban and rural communities and students of color who face structural barriers to STEM learning—may be required to use a scripted curriculum or shoulder additional responsibilities not imposed on teachers in wealthier and predominantly white schools. Addressing these disparities will not come through a singular focus on the design, development, and use of GenAI for STEM teaching. Rather, it is critical to understand the impacts of GenAI in a sociocultural context with an excavation of the cultural, historical, and political factors that influence STEM teaching and learning in the classroom and to address those factors ***at the correct scale***. Otherwise, we sustain the tradition of “tinkering toward utopia” (Tyack & Cuban, 1997) while existing inequalities become more entrenched. Such comprehensive efforts, encompassing both technological advancements and systemic reforms, will not only enhance our understanding of GenAI’s potential in STEM education but also fortify our ability to address deeply rooted inequalities. This holistic approach is essential to ensuring a more equitable distribution of benefits across all educational settings and to leveraging GenAI as a tool for reducing, rather than reinforcing, existing educational disparities.

References

Adelman, H. S. (1971). Learning problems. *Academic Therapy, 6*(3), 287–292. https://doi.org/10.1177/105345127100600310

Aggarwal-Schifellite, M. (2017, June 14). A history of the subtle sexism of home technology. *Pacific Standard*. https://psmag.com/news/a-history-of-the-subtle-sexism-of-home-technology

Agrawal, G., Kumarage, T., Alghamdi, Z., & Liu, H. (2024). Can knowledge graphs reduce hallucinations in LLMs? : A Survey (arXiv:2311.07914). *arXiv*. http://arxiv.org/abs/2311.07914

Al Naqbi, H., Bahroun, Z., & Ahmed, V. (2024). Enhancing work productivity through generative artificial intelligence: A comprehensive literature review. *Sustainability, 16*(3). https://doi.org/10.3390/su16031166

Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education, 100*(8), 2965–2971. https://doi.org/10.1021/acs.jchemed.3c00323

Alcaraz, A. (2024, May 21). Harnessing knowledge graphs to mitigate hallucinations in large language models. *Medium*. https://medium.com/codex/harnessing-knowledge-graphs-to-mitigate-hallucinations-in-large-language-models-d6fa6c7db07e

Aleven, V., Roll, I., McLaren, B. M., & Koedinger, K. R. (2016). Help helps, But only so much: Research on help seeking with intelligent tutoring systems. *International Journal of Artificial Intelligence in Education, 26*(1), 205–223. https://doi.org/10.1007/s40593-015-0089-1

Alim, H. S., Paris, D., & Wong, C. P. (2020). Culturally sustaining pedagogy: A critical framework for centering communities. In N. S. Nasir, C. D. Lee, R. Pea, & M. M. deRoyston (Eds.), *Handbook of the cultural foundations of learning* (pp. 261–276). Routledge.

Amador, J. M., Rogers, M. A. P., Hudson, R., Phillips, A., Carter, I., Galindo, E., & Akerson, V. L. (2022). Novice teachers' pedagogical content knowledge for planning and implementing mathematics and science lessons. *Teaching and Teacher Education, 115*, 103736. https://doi.org/10.1016/j.tate.2022.103736

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4*(2), 167–207. https://doi.org/10.1207/s15327809jls0402_2

Anderson, J. R., & Tully, D. (2024). Creating new STEM opportunities in schools: Teachers as curriculum designers. In Y. Li, Z. Zeng, & N. Song (Eds.), *Disciplinary and interdisciplinary education in STEM* (pp. 265–293). Springer, Cham. https://doi.org/10.1007/978-3-031-52924-5_13

Anthropic. (2024). System prompts. https://docs.anthropic.com/en/release-notes/system-prompts

Aronson, B., & Laughter, J. (2016). The theory and practice of culturally relevant education: A synthesis of research across content areas. *Review of Educational Research, 86*(1), 163–206. https://doi.org/10.3102/0034654315582066

Baker, R. S., & Rossi, L. M. (2013). Assessing the disengaged behaviors of learners. *Design Recommendations for Intelligent Tutoring Systems, 1*, 153.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6–14. https://doi.org/10.3102/0013189X025009006

Ballance, L., & Ripley, E. (2023). Paying research participants and community and patient research partners: An engaging ethical issue. In E. E. Anderson (Ed.), *Ethical issues in community and patient stakeholder–Engaged health research* (pp. 181–195). Springer International Publishing. https://doi.org/10.1007/978-3-031-40379-8_14

Bang, M., Faber, L., Gurneau, J., Marin, A., & Soto, C. (2016). Community-based design research: Learning across generations and strategic transformations of institutional relations toward axiological innovations. *Mind, Culture, and Activity, 23*(1), 28–41. https://doi.org/10.1080/10749039.2015.1087572

Bang, M., & Vossoughi, S. (2016). Participatory design research and educational justice: Studying learning and relations within social change making. *Cognition and Instruction, 34*(3), 173–193. https://doi.org/10.1080/07370008.2016.1181879

Barnes, T., Burriss, S., Danish, J., Finkelstein, S., Humburg, M., Limke, A., Molvig, O., & Reichert, H. (2024). *Toward ethical and just AI in education research*. Community for Advancing Discovery Research in Education (CADRE). https://eric.ed.gov/?id=ED655704

Barros, A. de. (2023). *Explaining the productivity paradox: Experimental evidence from educational technology* (EdWorkingPaper: 23-853). Annenberg Institute for School Reform at Brown University. https://eric.ed.gov/?id=ED639260

Ben-Peretz, M. (1975). The concept of curriculum potential. *Curriculum Theory Network, 5*(2), 151–159. https://www.tandfonline.com/doi/abs/10.1080/00784931.1975.11075803

Berberette, E., Hutchins, J., & Sadovnik, A. (2024). Redefining “hallucination” in LLMs: Toward a psychology-informed framework for mitigating misinformation (arXiv:2402.01769). *arXiv*. http://arxiv.org/abs/2402.01769

Blackhurst, A. E. (1965). Technology in special education–Some implications. *Exceptional Children, 31*(9), 449–456. https://doi.org/10.1177/001440296503100901

Blanchett, W. J., Klingner, J. K., & Harry, B. (2009). The intersection of race, culture, language, and disability: Implications for urban education. *Urban Education, 44*(4), 389–409. https://doi.org/10.1177/0042085909338686

Borko, H., Livingston, C., & Shavelson, R. J. (1990). Teachers' thinking about instruction. *Remedial and Special Education, 11*(6), 40–49. https://doi.org/10.1177/074193259001100609

Brown, B. A., Henderson, J. B., Gray, S., Donovan, B., Sullivan, S., Patterson, A., & Waggstaff, W. (2016). From description to explanation: An empirical exploration of the African-American pipeline problem in STEM. *Journal of Research in Science Teaching, 53*(1), 146–177. https://doi.org/10.1002/tea.21249

Brown, J., Castagno, A., Dass, P., Davidson, S., Jaber, L., Joseph, D. H., Kramarczuk, K. A., Ketelhut, D. J., Metcalf, A., & Pacheco, M. (2022, April 4). *Culturally responsive science instruction. Community for Advancing Discovery Research in Education*. https://cadrek12.org/events/2022/04/04/culturally-responsive-science-instruction

Buxton, C. A., & Allexsaht-Snider, M. (2017). *Supporting K-12 english language learners in science*. Routledge.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work* (Working paper 31161). National Bureau of Economic Research. https://doi.org/10.3386/w31161

Calo, R., Friedman, B., Kohno, T., Almeter, H., & Logler, N. (Eds.). (2021). *Telling stories: On culturally responsive artificial intelligence*. University of Washington Tech Policy Lab.

Candrian, C., & Scherer, A. (2023). Reactance to human versus artificial intelligence: Why positive and negative information from human and artificial agents leads to different responses. https://doi.org/10.2139/ssrn.4397618

Carroll, S. R., Garba, I., Plevel, R., Small-Rodriguez, D., Hiratsuka, V. Y., Hudson, M., & Garrison, N. A. (2022). Using Indigenous standards to implement the CARE principles: Setting expectations through tribal research codes. *Frontiers in Genetics, 13*, 823309. https://doi.org/10.3389/fgene.2022.823309

CASEL. (2024). What is the CASEL framework? https://casel.org/fundamentals-of-sel/what-is-the-casel-framework/

Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences, 108*(supplement_3), 15655–15659. https://doi.org/10.1073/pnas.1010658108

Su-Yen Chen. (2023). Generative AI, learning and new literacies: Journal of Educational Technology Development & Exchange. *Journal of Educational Technology Development & Exchange, 16*(2), 1–19. https://doi.org/10.18785/jetde.1602.01

Cherbow, K., & McNeill, K. L. (2022). Planning for student-driven discussions: A revelatory case of curricular sensemaking for epistemic agency. *Journal of the Learning Sciences, 31*(3), 408–457. https://doi.org/10.1080/10508406.2021.2024433

Chiu, J. L., Fick, S. J., McElhaney, K. W., Alozie, N., & Fujii, R. (2021). Elementary teacher adaptations to engineering curricula to leverage student and community resources. *Journal of Pre-College Engineering Education Research (J-PEER), 11*(1). https://doi.org/10.7771/2157-9288.1283

Chiu, T. K. F. (2023). The impact of generative ai (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 1–17. https://doi.org/10.1080/10494820.2023.2253861

Christie, P., & Lingard, B. (2020). *Teachers and schooling making a difference: Productive pedagogies, assessment and performance*. Routledge.

Cioè-Peña, M. (2022). TrUDL, A path to full inclusion: The intersectional possibilities of translanguaging and universal design for learning. *TESOL Quarterly, 56*(2), 799–812. https://doi.org/10.1002/tesq.3074

Cohn, C., Snyder, C., Montenegro, J., & Biswas, G. (2024). Toward a human-in-the-loop LLM approach to collaborative discourse analysis (Vol. 2151, pp. 11–19). https://doi.org/10.1007/978-3-031-64312-5_2

Colace, F., Lombardi, M., & Santaniello, D. (2020). A Chatbot for supporting users in cultural heritage contexts (p. 27). https://doi.org/10.18293/DMSVIVA20-009

Cook-Sather, A. (2020). Student voice across contexts: Fostering student agency in today's schools. *Theory Into Practice, 59*(2), 182–191. https://doi.org/10.1080/00405841.2019.1705091

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology, 32*(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y

Copur-Gencturk, Y., Li, J., Cohen, A. S., & Orrill, C. H. (2024). The impact of an interactive, personalized computer-based teacher professional development program on student performance: A randomized controlled trial. *Computers & Education, 210*, 104963. https://doi.org/10.1016/j.compedu.2023.104963

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media, 20*(4), 336–349. https://doi.org/10.1177/1527476418796632

Couldwell, D. (2024). Overcoming AI hallucinations with RAG and knowledge graphs. *InfoWorld*. https://www.infoworld.com/article/3511437/overcoming-ai-hallucinations-with-rag-and-knowledge-graphs.html

Crawford, K. (2024). Generative AI's environmental costs are soaring—And mostly secret. *Nature, 626*(8000), 693–693. https://doi.org/10.1038/d41586-024-00478-x

de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford.

De Cremer, D., & Kasparov, G. (2021). AI should augment human intelligence, not replace it. *Harvard Business Review, 18*.

Dede, C., Eisenkraft, A., Frumin, K., & Hartley, A. (2016). *Teacher learning in the digital age: Online professional development in STEM education*. Harvard Education Press.

Demszky, D., & Liu, J. (2023). M-powering teachers: Natural language processing powered feedback improves 1:1 instruction and student outcomes. *Proceedings of the Tenth ACM Conference on Learning @ Scale*, 59–69. https://doi.org/10.1145/3573051.3593379

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding (arXiv:1810.04805). *arXiv.* https://doi.org/10.48550/arXiv.1810.04805

Diliberti, M., Schwartz, H. L., Doan, S., Shapiro, A. K., Rainey, L., & Lake, R. J. (2024). *Using artificial intelligence tools in K–12 classrooms*. RAND. https://www.rand.org/content/dam/rand/pubs/research_reports/RRA900/RRA956-21/RAND_RRA956-21.pdf

Edouard, K. (2024). Black prompts not included: AI bias impeding real-time co-creative ideation in black students' conceptual design. *Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024* (pp. 1366–1369). International Society of the Learning Sciences. https://doi.org/10.22318/icls2024.755248

Emara, M., Hutchins, N. M., Grover, S., Snyder, C., & Biswas, G. (2021). Examining student regulation of collaborative, computational, problem-solving processes in open-ended learning environments. *Journal of Learning Analytics, 8*(1), 49–74. https://doi.org/10.18608/jla.2021.7230

Eriksen, A. S., & Eriksen, A. P. (2023). Synthesizing minds and machines: An empirical study of the impact of human-ai co-creation on creativity and the moderating role of competence [Master thesis, Norwegian School of Economics]. https://openaccess.nhh.no/nhh-xmlui/handle/11250/3129712

Feng, M., Heffernan, N., Collins, K., Heffernan, C., & Murphy, R. F. (2023). Implementing and evaluating ASSISTments online math homework support at large scale over two years: Findings and lessons learned. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova (Eds.), *Artificial intelligence in education* (Vol. 13916, pp. 28–40). Springer Champ. https://doi.org/10.1007/978-3-031-36272-9_3

Fischer, C. T., & Rizzo, A. A. (1974). A paradigm for humanizing special education. *Journal of Special Education, 8*(4), 321–329. https://doi.org/10.1177/002246697400800404

Fritzgerald, A. (2020). Antiracism and universal design for learning. CAST.

Gage, M., Pizer, A., & Roth, V. (2001). WeBWorK: An internet-based system for generating and delivering homework problems. Joint Meeting of the American Mathematical Society and the Mathematical Association of America.

Gagné, R. M. (1974). Educational technology and the learning process. *Educational Researcher, 3*(1), 3–8. https://doi.org/10.3102/0013189X003001004

Gardner, H. (2009, May/June). Personalized education. *Foreign Policy, 172*, 86. https://www.proquest.com/openview/fe568ed046857e363c1d72797a913499/1?pq-origsite=gscholar&cbl=47510

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Garcia Coll, C., Garcia Miranda, A., Buzzetta Torres, I., & Nogueras Bermúdez, J. (2018). On becoming cultural beings: A focus on race, gender, and language. *Research in Human Development, 15*(3–4), 332–344. https://doi.org/10.1080/15427609.2018.1491217

Gobert, J. D., Sao Pedro, M. A., & Betts, C. G. (2023). An AI-based teacher dashboard to support students' inquiry: Design principles, features, and technological specifications. In *Handbook of research on science education*, Vol. III (pp. 1011–1044). Routledge.

Gobert, J. D., Sao Pedro, M., Raziuddin, J., & Baker, R. S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences, 22*(4), 521–563. https://doi.org/10.1080/10508406.2013.837391

Gould, M. A. (2008). Teacher as researcher: A paradigm for professional development. *Kappa Delta Pi Record, 45*(1), 5–7. https://doi.org/10.1080/00228958.2008.10516523

Griscom, R. (2024). Bill Gates says superhuman AI may be closer than you think. *Next Big Idea Club*. https://nextbigideaclub.com/magazine/bill-gates-says-superhuman-ai-may-closer-think-podcast/50267/

Grover, S., Basu, S., Bienkowski, M., Eagle, M., Diana, N., & Stamper, J. (2017). A framework for using hypothesis-driven approaches to support data-driven learning analytics in measuring computational thinking in block-based programming environments. *ACM Transactions on Computer Education, 17*(3), 1–25. https://doi.org/10.1145/3105910

Grover, S., & Clarkson, C. (2023). Empowering math teachers with an AI tool for auto-generation of technology-enhanced assessments. STELAR. https://stelar.edc.org/projects/24644/profile/empowering-math-teachers-ai-tool-auto-generation-technology-enhanced

Grover, S., & Korhonen, A. (2017). Unlocking the potential of learning analytics in computing education. *ACM Transactions on Computer Education, 17*(3), 1–4. https://doi.org/10.1145/3122773

Grover, S., Runge, B., Lynch, R., & Clarkson, C. (2024). Teachers' use of a generativeAI assessment creation tool in high school mathematics classrooms. In *Proceedings of the 12th International Conference on Technology 4 Education*. London, UK. Springer Nature.

Guo, Q. (2024). Prompting change: ChatGPT's impact on digital humanities pedagogy – A case study in art history. *International Journal of Humanities and Arts Computing, 18*(1), 58–78. https://doi.org/10.3366/ijhac.2024.0321

Hajli, M., Sims, J. M., & Ibragimov, V. (2015). Information technology (IT) productivity paradox in the 21st century. *International Journal of Productivity and Performance Management, 64*(4), 457–478. https://doi.org/10.1108/IJPPM-12-2012-0129

Harris, C. J., Wiebe, E., Grover, S., & Pellegrino, J. W. (2023). *Classroom-Based STEM assessment: Contemporary issues and perspectives*. Community for Advancing Discovery Research in Education (CADRE). https://eric.ed.gov/?id=ED631656

Harrison, R. (1978). How to design and conduct self-directed learning experiences. *Group & Organization Studies, 3*(2), 149–167. https://doi.org/10.1177/105960117800300203

Hayes, D. (2003). Making learning an effect of schooling: Aligning curriculum, assessment and pedagogy. *Discourse: Studies in the Cultural Politics of Education, 24*(2), 225–245. https://doi.org/10.1080/01596300303039

Hendry, D. G., Friedman, B., & Ballard, S. (2021). Value sensitive design as a formative framework. *Ethics and Information Technology, 23*(1), 39–44. https://doi.org/10.1007/s10676-021-09579-x

Hennessy, S., Deaney, R., & Ruthven, K. (2005). Emerging teacher strategies for mediating 'Technology-integrated Instructional Conversations': A socio-cultural perspective. *The Curriculum Journal, 16*(3), 265–292. https://doi.org/10.1080/09585170500256487

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570.

Hopkins, M., Spillane, J. P., Jakopovic, P., & Heaton, R. M. (2013). Infrastructure redesign and instructional reform in mathematics: Formal structure and teacher leadership. T*he Elementary School Journal, 114*(2), 200–224. https://doi.org/10.1086/671935

Hopkins, M., & Woulfin, S. L. (2015). School system (re) design: Developing educational infrastructures to support school leadership and teaching practice. *Journal of Educational Change, 16*(4), 371–377.

Hu, W., & Guo, X. (2021). Toward the development of key competencies: A conceptual framework for the STEM curriculum design and a case study. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.684265

Huang, B., Jong, M. S.-Y., Tu, Y.-F., Hwang, G.-J., Chai, C. S., & Jiang, M. Y.-C. (2022). Trends and exemplary practices of STEM teacher professional development programs in K–12 contexts: A systematic review of empirical studies. *Computers & Education, 189*, 104577. https://doi.org/10.1016/j.compedu.2022.104577

Hutchins, N. M., & Biswas, G. (2023a). Catalyzing teachers' evidence-based responses to students' problem-based learning in STEM. ISLS Annual Meeting, Montreal, Quebec, CA. https://doi.org/10.22318/icls2023.131431

Hutchins, N. M., & Biswas, G. (2023b). Using teacher dashboards to customize lesson plans for a problem-based, middle school STEM curriculum (pp. 324–332). LAK23: 13th International Learning Analytics and Knowledge Conference, Arlington, TX, US. https://doi.org/10.1145/3576050.3576100

Hutchins, N. M., & Biswas, G. (2024). Co-designing teacher support technology for problem-based learning in middle school science. *British Journal of Educational Technology, 55*(3), 802–822. https://doi.org/10.1111/bjet.13363

Ikpeze, C. H. (2016). *Teaching as a cultural artifact*. Brill.

IVOW. (2018, April 24). Can we train the next generation of #AI storytellers to be culturally conscious? *Medium*. https://medium.com/@IVOW.ai/can-we-train-the-next-generation-of-ai-storytellers-to-be-culturally-conscious-a619b505a139

IVOW. (2020, June 4). *Bringing cultural context to artificial intelligence*. https://issuu.com/ivowai/docs/whitepaper_ikg_ivow

IVOW. (2021). *Indigenous knowledge graph*. https://issuu.com/ivowai/docs/ivow_learnings_final/s/15950414

Jenkins, J. M. (1998). Properties of personalized learning. *International Journal of Educational Reform, 7*(1), 66–71. https://doi.org/10.1177/105678799800700107

Johnson, E. (2024, July 31). 'Intersectional hallucinations': Why AI struggles to understand that a six-year-old can't be a doctor or claim a pension. *The Conversation*. http://theconversation.com/intersectional-hallucinations-why-ai-struggles-to-understand-that-a-six-year-old-cant-be-a-doctor-or-claim-a-pension-234879

Johnson, K. (2023). Teachers are going all in on generative AI. *Wired.* https://www.wired.com/story/teachers-are-going-all-in-on-generative-ai/

Johnson, K. (2024, June 3). California teachers are using AI to grade papers. Who's grading the AI? *CalMatters.* http://calmatters.org/economy/technology/2024/06/teachers-ai-grading/

Kaggwa, R. J., Wester, E., Arango-Caro, S., Woodford-Thomas, T. A., & Callis-Duehl, K. (2023). STEM outreach to underresourced schools: A model for inclusive student engagement. *Journal of STEM Outreach, 6*(1). https://doi.org/10.15695/jstem/v6i1.04

Karpouzis, K., Pantazatos, D., Taouki, J., & Meli, K. (2024). Tailoring education with GenAI: A new horizon in lesson planning (arXiv:2403.12071). *arXiv*. https://doi.org/10.48550/arXiv.2403.12071

Keratithamkul, K., Kim, J. N., & Roehrig, G. H. (2020). Cultural competence or deficit-based view? A qualitative approach to understanding middle school students' experience with culturally framed engineering. *International Journal of STEM Education, 7*(1), 1–15. https://doi.org/10.1186/S40594-020-00224-5

Klassen, S., & Fiesler, C. (2022). "This isn't your data, friend": Black Twitter as a case study on research ethics for public data. *Social Media + Society, 8*(4). https://doi.org/10.1177/20563051221144317

Ko, M.-L. M., Superfine, A. C., & Goldman, S. R. (2022). Interacting and intersecting contexts of teacher learning: Next steps for learning sciences research. In A. C. Superfine, S. R. Goldman, & M-L. M. Ko (Eds.), *Teacher learning in changing contexts*. Routledge.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30–43.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36*(5), 757–798. https://doi.org/10.1111/j.1551-6709.2012.01245.x

Kulik, J. A., Kulik, C. C., & Smith, B. B. (1976). Research on the personalized system of instruction. *Programmed Learning and Educational Technology, 13*(1), 23–30. https://doi.org/10.1080/1355800760130104

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Ladson-Billings, G. (2021). Three decades of culturally relevant, responsive, & sustaining pedagogy: What lies ahead? *The Educational Forum, 85*(4), 351–354. https://doi.org/10.1080/00131725.2021.1957632

Langran, E., Searson, M., & Trumble, J. (2024). Transforming teacher education in the age of generative AI. In M. Searson, E. Langran, & J. Trumble (Eds.), *Exploring new horizons: generative artificial intelligence and teacher education* (pp. 2–13). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/p/223928/

Laster, B., Shargel, R., & Watson-Vandiver, M. (2020). Culturally sustaining pedagogy extended: Ethnic pedagogies. *Multicultural Perspectives, 22*(2), 91–100. https://doi.org/10.1080/15210960.2020.1741366

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Learning for Justice. (2023, May 23). Critical practices for social justice education. *Learning for Justice*. https://www.learningforjustice.org/magazine/publications/critical-practices-for-social-justice-education

Lee, H.-S., Gweon, G.-H., Lord, T., Paessel, N., Pallant, A., & Pryputniewicz, S. (2021). Machine learning-enabled automated feedback: Supporting students' revision of scientific arguments based on data drawn from simulation. *Journal of Science Education and Technology, 30*(2), 168–192. https://doi.org/10.1007/s10956-020-09889-7

Lee, O. (2021). Asset-oriented framing of science and language learning with multilingual learners. *Journal of Research in Science Teaching, 58*(7), 1073–1079. https://doi.org/10.1002/tea.21694

Lee, S. (2012, Fall). Knowing myself to know others: Preparing preservice teachers for diversity through multicultural autobiography. *Multicultural Education*, 38–41.

Lee, V. R., Pope, D., Miles, S., & Zárate, R. C. (2024, December). 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*, 100253. https://doi.org/10.1016/j.caeai.2024.100253

Lertdechapat, K., & Faikhamta, C. (2021). Enhancing pedagogical content knowledge for STEM teaching of teacher candidates through lesson study. *International Journal for Lesson & Learning Studies, 10*(4), 331–347. https://doi.org/10.1108/IJLLS-03-2021-0020

Liu, S., Yu, Z., Huang, F., Bulbulia, Y., Bergen, A., & Liut, M. (2024, July 3). Can small language models with retrieval-augmented generation replace large language models when learning computer science? *ITiCSE: 2024 Proceedings of the 2024 on Innovation and Technology in Computer Science Education* V. 1, 388–393. https://doi.org/10.1145/3649217.3653554

Luckin, R. (2025, January-a). *What the research says about AI tutoring* [Post]. LinkedIn. https://www.linkedin.com/posts/rose-luckin-5245003_skinnyonaied-ai-edtech-activity-7281950164280348672-BonY

Luckin, R. (2025, January-b). *What the research says: Step-based learning - a critical lesson for GenAI tutoring* [Post]. LinkedIn. https://www.linkedin.com/posts/rose-luckin-5245003_aied-aitutoring-learningscience-activity-7285211648938704896-B9Jy

Macfadyen, L. P. (2022). Institutional implementation of learning analytics: current state, Challenges, and guiding frameworks. In Lang, C., Siemens, G., Wise, A. F., Gašević, D., & Merceron, A. (Eds.), *The handbook of learning analytics, 2*, 173-186. Society for Learning Analytics Research.

Maier, U., & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. *Computers and Education: Artificial Intelligence, 3*, 100080. https://doi.org/10.1016/j.caeai.2022.100080

Malik, A. A., Hassan, M., Rizwan, M., Mushtaque, I., Lak, T. A., & Hussain, M. (2023). Impact of academic cheating and perceived online learning effectiveness on academic performance during the COVID-19 pandemic among Pakistani students. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1124095

Marco-Bujosa, L. M., McNeill, K. L., & Friedman, A. A. (2020). Becoming an urban science teacher: How beginning teachers negotiate contradictory school contexts. *Journal of Research in Science Teaching, 57*(1), 3–32. https://doi.org/10.1002/tea.21583

McFadden, J. R., & Roehrig, G. H. (2017). Exploring teacher design team endeavors while creating an elementary-focused STEM-integrated curriculum. *International Journal of STEM Education, 4*(1), 21. https://doi.org/10.1186/s40594-017-0084-1

McKee, K. R. (2024). Human participants in AI research: Ethics and transparency in practice. *IEEE Transactions on Technology and Society, 5*(3), 279–288. https://doi.org/10.1109/TTS.2024.3446183

Midjaas, C. L. (1970). The middle school: An opportunity for humanized education. https://eric.ed.gov/?id=ED046110

Miller, E. C., Severance, S., & Krajcik, J. (2021). Motivating teaching, sustaining change in practice: Design principles for teacher learning in project-based learning contexts. *Journal of Science Teacher Education, 32*(7), 757–779. https://doi.org/10.1080/1046560X.2020.1864099

Muller, M., Weisz, J. D., & Geyer, W. (n.d.). Mixed initiative generative AI interfaces: An analytic framework for generative AI applications. IBM Research – Cambridge. https://computationalcreativity.net/workshops/cocreative-iccc20/papers/Future_of_co-creative_systems_185.pdf

National Academies of Sciences, Engineering, and Medicine. (2022). *Human-AI teaming: State-of-the-art and research needs*. National Academies Press. https://doi.org/10.17226/26355

National Research Council. (2011). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. https://doi.org/10.17226/13165

National Science Foundation. (2023, December 1). Advancing education for the future AI workforce (EducateAI) https://new.nsf.gov/funding/opportunities/advancing-education-future-ai-workforce-educateai

Nespor, J. (1997). *Tangled up in school: Politics, space, bodies, and signs in the educational process*. Lawrence Erlbaum.

Nieto, S. (2008). Culture and education. *Teachers College Record: The Voice of Scholarship in Education, 110*(13), 127–142. https://doi.org/10.1177/016146810811001311

Noddings, N. (1988). An ethic of caring and its implications for instructional arrangements. *American Journal of Education, 96*(2), 215–230. https://doi.org/10.1086/443894

Noddings, N. (2010). Moral education in an age of globalization. *Educational Philosophy and Theory, 42*(4), 390–396. https://doi.org/10.1111/j.1469-5812.2008.00487.x

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

NYSED.gov. (2018). *Culturally responsive-sustaining education framework*. New York State Department of Education. https://www.nysed.gov/sites/default/files/programs/crs/culturally-responsive-sustaining-education-framework.pdf

Ocumpaugh, J., Roscoe, R. D., Baker, R. S., Hutt, S., & Aguilar, S. J. (2024). Toward asset-based instruction and assessment in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-023-00382-x

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher, 41*(3), 93–97. https://doi.org/10.3102/0013189X12441244

Paris, D., & Alim, H. S. (Eds.). (2017). *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. Teachers College Press.

Penuel, W. R. (2019). Infrastructuring as a practice of design-based research for supporting and studying equitable implementation and sustainability of innovations. *Journal of the Learning Sciences, 28*(4–5), 659–677. https://doi.org/10.1080/10508406.2018.1552151

Perrigo, B. (2023, January 18). Exclusive: The $2 per hour workers who made ChatGPT safer. *TIME*. https://time.com/6247678/openai-chatgpt-kenya-workers/

Postman, N. (1998). Five things we need to know about technological change. https://www.cs.ucdavis.edu/~rogaway/classes/188/materials/postman.pdf

Price, J. F. (2024, June 24). Unpacking the AI mirror: Illuminating biases in Chatbots' stories of learning. ISTE Live 2024, Denver, CO, USA. https://conference.iste.org/2024/program/search/detail_session.php?id=117202962

Price, J. F., Chakraborty, S., Hall, T., Waechter-Versaw, A., & Santamaría Graff, C. (2023). RAPID: DRL AI: A community-inclusive AI chatbot to support teachers in developing culturally focused and universally designed STEM activities [Award abstract 2334631]. National Science Foundation. https://www.nsf.gov/awardsearch/showAward?AWD_ID=2334631&HistoricalAwards=false

Price, J. F., & McNeill, K. L. (2013). Toward a lived science curriculum in intersecting figured worlds: An exploration of individual meanings in science education. *Journal of Research in Science Teaching, 50*(5), 501–529. https://doi.org/10.1002/tea.21084

Priestley, M., Biesta, G. J. J., Philippou, S., & Robinson, S. (2015). The teacher and the curriculum: Exploring teacher agency. In D. Wyse, L. Hayward, & J. Pandya (Eds.), *The SAGE handbook of curriculum, pedagogy and assessment* (pp.187–201). Sage.

Rabinowitz, M., González, I., & Rogers, R. C. (2024). The equity-centered participatory compensation model (EPCM): A tutorial for this emergent methodology. *Journal of Participatory Research Methods, 5*(1). https://doi.org/10.35844/001c.92362

Rathle, P. (2024, July 11). The GraphRAG manifesto: Adding knowledge to GenAI. *Neo4j*. https://neo4j.com/blog/graphrag-manifesto/

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246. https://doi.org/10.3102/00346543075002211

Rose, D., & Meyer, A. (2000). The future is in the margins: The role of technology and disability in educational reform. http://eric.ed.gov/?id=ED451624%20

Rose, D., & Meyer, A. (2002). *Teaching every student in the digital age: Universal design for learning*. Association for Supervision and Curriculum Development.

Sakamoto, M., Tan, S., & Clivaz, S. (2024). Social, cultural and political perspectives of generative AI in teacher education: Lesson planning in Japanese teacher education. In M. Searson, E. Langran, & J. Trumble (Eds.). *Exploring new horizons: Generative artificial intelligence and teacher education*. Association for the Advancement of Computing in Education.

Scharff, D. P., Mathews, K. J., Jackson, P., Hoffsuemmer, J., Martin, E., & Edwards, D. (2010). More than Tuskegee: Understanding mistrust about research participation. *Journal of Health Care for the Poor and Underserved, 21*(3), 879–897. https://doi.org/10.1353/hpu.0.0323

Schwab, J. J. (1973). The practical 3: Translation into curriculum. *The School Review, 81*(4), 501–522.

Shute, V. J., & Ventura, M. (2015). Stealth assessment. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology* (pp. 675–676). Sage.

Sims, S., & Fletcher-Wood, H. (2021). Identifying the characteristics of effective teacher professional development: A critical review. *School Effectiveness and School Improvement, 32*(1), 47–63. https://doi.org/10.1080/09243453.2020.1772841

Slota, S. C., Yi, S., Fleischmann, K. R., Bailey, J., & Watkins, S. C. (2021, November). Methods for eliciting feedback about AI and racial equity. Technology, Mind, and Society, El Paso, TX, USA.

Solorzano, D. G., & Yosso, T. J. (2001). From racial stereotyping and deficit discourse toward a critical race theory in teacher education. *Multicultural Education; San Francisco, 9*(1), 2–8.

Stahl, L., Chasan, A., Bar-On, S., & Jung, J. (2024, November 24). Kenyan workers with AI jobs thought they had tickets to the future until the grim reality set in. *CBS News*. https://www.cbsnews.com/news/ai-work-kenya-exploitation-60-minutes/

Stanford Neurodiversity Project. (2024). *Neurodiversity at work*. Stanford Neurodiversity Project. https://med.stanford.edu/neurodiversity/NaW.html

Stephenson Reaves, J. R., Likely, R., & Arias, A. M. (2022). Design principles for considering the participatory relationship of students, teachers, curriculum, and place in project-based STEM units. *Education Sciences, 12*(11), 760. https://doi.org/10.3390/educsci12110760

Sun, Y., Jang, E., Ma, F., & Wang, T. (2024). Generative AI in the wild: Prospects, challenges, and strategies. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3613904.3642160

Thatcher, J., O'Sullivan, D., & Mahmoudi, D. (2016). Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space, 34*(6), 990–1006. https://doi.org/10.1177/0263775816633195

Treleaven, P., & Brown, D. (2024). *AI for social good* (SSRN scholarly paper 4746471). Social Science Research Network. https://doi.org/10.2139/ssrn.4746471

Trump, J. L. (1961). Basic changes needed to serve individuals better. *The Educational Forum, 26*(1), 93–101. https://doi.org/10.1080/00131726109340304

Tyack, D. B., & Cuban, L. (1997). *Tinkering toward utopia: A century of public school reform*. Harvard University Press.

U.S. Department of Education, Office of Educational Technology. (2023). *Artificial intelligence and the future of teaching and learning*. https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf

U.S. Department of Education, Office of Educational Technology. (2024). *Designing for education with artificial intelligence: An essential guide for developers*. https://tech.ed.gov/files/2024/07/Designing-for-Education-with-Artificial-Intelligence-An-Essential-Guide-for-Developers.pdf

Vallor, S. (2024). *The AI mirror: How to reclaim our humanity in an age of machine thinking*. Oxford.

Vaughn, M. (2020). What is student agency and why is it needed now more than ever? *Theory Into Practice, 59*(2), 109–118. https://doi.org/10.1080/00405841.2019.1702393

Wang, H., Haudek, K. C., Manzanares, A. D., Romulo, C. L., & Royse, E. A. (2024). Extending a pretrained language model (BERT) using an ontological perspective to classify students' scientific expertise level from written responses. https://doi.org/10.21203/rs.3.rs-3879583/v1

Wang, X. (2023). Empowering teachers to collaborate with generative AI for developing high-quality STEM learning resources. STELAR – STEM Learning and Research Center. https://stelar.edc.org/projects/24657/profile/empowering-teachers-collaborate-generative-ai-developing-high-quality-stem

Watkins, S. C. (2021). *Artificial intelligence and the future of racial justice*. TED Talks. https://www.ted.com/talks/s_craig_watkins_artificial_intelligence_and_the_future_of_racial_justice

Wignall, G., & Hart, E. (2024, August 26). School-research partnerships – plagiarism and student use of AI tools. *Teacher Magazine*. https://www.teachermagazine.com/au_en/articles/school-research-partnerships-plagiarism-and-student-use-of-ai-tools

Williamson, K., & Kizilcec, R. (2022, March 21). A review of learning analytics dashboard research in higher education: Implications for justice, equity, diversity, and inclusion (pp. 260–270). LAK22: 12th International Learning Analytics and Knowledge Conference. ACM. https://doi.org/10.1145/3506860.3506900

Wolfson, B. J. (1968). Pupil and teacher roles in individualized instruction. *Elementary School Journal, 68*(7), 357–366. https://doi.org/10.1086/460459

Yin, Y., Jia, N., & Wakslak, C. J. (2024). AI can help people feel heard, but an AI label diminishes this impact. *Proceedings of the National Academy of Sciences, 121*(14), e2319112121. https://doi.org/10.1073/pnas.2319112121

Yosso, T. J. (2005). Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education, 8*(1), 69–91. https://doi.org/10.1080/1361332052000341006

Zhai, X., Yin, Y., Pellegrino, J. W., Haudek, K. C., & Shi, L. (2020). Applying machine learning in science assessment: A systematic review. *Studies in Science Education, 56*(1), 111–151. https://doi.org/10.1080/03057267.2020.1735757

Zhou, M., Abhishek, V., Derdenger, T., Kim, J., & Srinivasan, K. (2024). Bias in generative AI (arXiv:2403.02726). *arXiv*. https://doi.org/10.48550/arXiv.2403.02726

Ziakkas, D., & Pechlivanis, K. (2023). Artificial intelligence applications in aviation accident classification: A preliminary exploratory study. *Decision Analytics Journal, 9*, 100358. https://doi.org/10.1016/j.dajour.2023.100358

Community for Advancing Discovery Research in Education

CADRE is a network for STEM education researchers funded by the National Science Foundation's Discovery Research PreK-12 (DRK-12) program. Through in-person meetings, a website, common interest groups, newsletters, and more, CADRE connects these researchers who are endeavoring to improve education in science, technology, engineering, and mathematics in, and outside of, our schools.

CADRE helps DRK-12 researchers share their methods, findings, results, and products inside the research and development community and with the greater public so that we are:

- **Better informed** about the work that is being done,
- **Continually building** on what we have collectively learned,
- **Working with our schools, communities, and policy-makers** to make our findings and products accessible and usable, and
- **Progressively able to address new and more challenging issues—**including those issues that extend beyond the limits of what any singular research project can impact.

Together, we can make a larger impact on policy, research, and education.

Contact cadre@edc.org or visit cadrek12.org for more information.

EDC.ORG