C1 Misinfo

Gen AI has uniquely opened the door to misinfo

Welle 24 Deutsche Welle, 3-26-2024, "Generative AI is the ultimate disinformation amplifier", https://akademie.dw.com/en/generative-ai-is-the-ultimate-disinformation-amplifier/a-68593890

Generative artificial intelligence (GAI) adds a new dimension to the problem of disinformation. Freely available and largely unregulated tools make it possible for anyone to generate false information and fake content in vast quantities. These include imitating the voices of real people and creating photos and videos that are indistinguishable from real ones. But there is also a positive side. Used smartly, GAI can provide a greater number of content consumers with trustworthy information, thereby counteracting disinformation. To understand the positives and negatives of GAI, it is first important to understand what AI is, and what is so special about generative AI. **What do machine learning, AI and generative AI mean?** Artificial intelligence refers to a collection of ideas, technologies and techniques that relate to a computer system's capacity to perform tasks that normally require human intelligence. When we talk about AI in the context of journalism, we usually mean machine learning (ML) as a sub field of AI. In basic terms, machine learning is the process of training a piece of software, called a model, to make useful predictions or generate content from data. The roots of machine learning are in statistics, which can also be thought of as the art of extracting knowledge from data. What machine learning does is to use data to answer questions. More formally, it refers to the use of algorithms that learn patterns from data and can perform tasks without being explicitly programmed to do so. Or in other words: they learn. A language model (LM) is a machine learning model that aims to predict and generate plausible language (natural or human-like language). To put it very simply, it's basically a probability model that, using a data set and algorithm, predicts the next word in a sentence based on previous words. Such models are called generative models or generative AI, because they create new and original content and data. Traditional AI, on the other hand, focuses on performing preset tasks using preset algorithms, but doesn't create new content. When models are trained on enormous amounts of data, their complexity and efficacy increase. Early language models could predict the probability of a single word whereas modern large language models (LLMs) can predict the probability of sentences, paragraphs or even entire documents based on patterns used in the past. A key development in language modeling was the introduction in 2017 of Transformers, a deep learning architecture designed around the idea of attention mechanisms. This innovation allows the model to selectively focus on the most important part of the input for making the prediction, boosting a model's ability to capture crucial information. The computer science portal Geeks for Geeks gives Google Streetview's house number identification as an example of an attention mechanism in computer vision that enables models to systematically identify certain portions of an image for processing. Attention mechanisms also made it possible to process longer sequences by solving memory issues encountered in earlier models. Transformers are the state-of-the-art architecture for a wide variety of language model applications, such as translators and chatbots. ChatGPT, the best known chatbot, is based on a language model developed by OpenAI. It is built on the GPT (Generative Pre-trained Transformer) model architecture, and it is known for its natural language processing capabilities. **What does generative AI mean for disinformation?** Generative AI is the first technology to enter an area that was previously reserved for humans: the autonomous production of content in any form, and the understanding and creation of language and meaning. And this is precisely what links generative AI to the topic of disinformation — the fact that**, today, it is often impossible to tell if content originates from a human or a machine, and if we can trust what we read, see or hear.** Media users are beginning to understand that something is broken in their relation to media and are confused. **"Some of the indicators that we have historically used to decide we should trust a piece of information have become distorted,"** Vinton G. Cerf, known as one of the "fathers of the internet," said in a 2024 video podcast by the international law firm Freshfields Bruckhaus Deringer. **What are the risks of ChatGPT and open-source large language models?** Although generative AI tools are still unavailable in some countries because of their internet censorship laws and regulations, the launch of ChatGPT by OpenAI in November 2022 (and later on its alternatives) was a turning point. Now, a large part of the world's internet users have access to these powerful tools and can use them according to their own purposes — whether positive or negative. It also means that through widespread use, the models can continue to learn and become better and even more powerful. **But the underlying LLM used by ChatGPT and Google's Gemini (formerly Bard) are owned by their companies, that is they are proprietary models. This raises concerns about LLMs' lack of transparency, the use of personal data for training purposes and limited accessibility.** There's also significant debate on the ability to use chatbots to produce disinformation and fake content. While these two chatbots in particular have garnered significant attention, other powerful open-source large language models, the foundational technology behind these chatbots, are freely available. Research by Democracy Reporting International, a Berlin-based organization promoting democracy, found these open-source LLMs, when managed by someone with the relevant coding skills, can rival the quality of products like ChatGPT and Gemini. But, it warned in its December 2023 report, "[u]nlike their more prominent counterparts, ... **these LLMs frequently lack integrated safeguards, rendering them more susceptible to misuse** in the creation of misinformation or hate speech."** What concrete negative effects does GAI have on disinformation?** We are seeing a whole range of different disinformation created by GAI, from fully AI generated fake news websites to fake Joe Biden robocalls telling Democrats not to vote.

And with the technology developing so quickly, media systems are having trouble adapting to it, learning how to use it safely and preventing dangers, while researchers are scrambling to identify and analyze the impacts. From the user's point of view, **generative AI is causing a general loss of trust in the media and difficulties in verifying the truthfulness of content, especially around elections. Deep fakes can be used to create non-consensual explicit content using someone's likeness, leading to severe privacy violations and harm to individuals, particularly women and marginalized communities.** Problem 1: Volume, automation and amplification **With GAI, the volume of disinformation potentially becomes infinite rendering fact checking an insufficient tool. As the marginal costs of the production of disinformation fall towards zero, the costs of dissemination are also nearly zero thanks to social media.** On top of this, individuals can now use user-friendly apps to easily and quickly generate sophisticated and convincing GAI content such as deep fake videos and voice clones – content that previously needed entire teams of tech-savvy individuals to produce. **This democratization of deep fake technology lowers the barrier of entry for creating and disseminating false narratives and misleading content online. Malign actors can easily leverage chatbots to spread falsehood across the internet at record speed, regardless of the languag**e. Text-to-text chatbots, such as ChatGPT or Gemini, or image generators, such as Midjourney, DALL-E or Stable Diffusion, can be used to create massive amounts of text as well as highly realistic fake audio, images and videos to spread misinformation and disinformation. This can lead to false narratives, country-specific misinformation, manipulation of public opinion and even harm to individuals or organizations. In a 2023 study, researchers at the University of Zurich in Switzerland found that generative AI can produce accurate information that is easier to understand, but it can also produce more compelling disinformation. Participants also failed to distinguish between posts on X, formerly Twitter, written by GPT-3 and written by real people. GAI applications can be combined to automate the whole process of content production, distribution and amplification. Fully synthetic visual material can be produced from a text prompt, and websites can be programmed automatically. Problem 2: Disinformation and the public arena's structural transformation Digitization has been transforming the public sphere for some time now. Generative AI is yet another element fueling this transformation, but it shouldn't be viewed in isolation, with structural shifts mainly happening because of digital media, economic pressures on traditional media organizations and the reconfiguration of attention allocation and information flows. The increase in the volume of AI-generated content, coupled with the difficulty in recognizing that content is AI-generated, is an additional factor in the public sphere's transformation. Information pollution has more than one cause apart from deliberately generated disinformation. Emily M. Bender, a linguistics professor at the University of Washington, addressed this problem in testimony before the US House Committee on Science, Space and Technology. Issues *Some reputable media houses are quietly posting synthetic text as if it were real reporting* (venerable tech outlet CNET was one of them, although it says it has paused this for now after an outcry). But the content can be biased or inaccurate if algorithms aren't designed properly, or if the training data sets are inherently biased. *GAI can hallucinate*. That means, it can produce content that isn't based on existing data or examples provided during the training process but rather made up. In one infamous example, in its very first demonstration, Google's Bard chatbot (as Gemini was called at the time) claimed that the James Webb Space Telescope had captured the first images of a planet outside our solar system, which wasn't factually true. *GAI has turbocharged plagiarism*. NewsGuard from the Journalism Trust Initiative was thefirst to identify the emergence of AI content farms using AI to copy and rewrite content from mainstream sources without credit. NewsGuard has identified hundreds of additional unreliable AI-generated websites. *Trust in democratic processes and institutions is eroding*. **The more polluted our information ecosystem becomes with synthetic text, the harder it will be to find trustworthy sources of information,** and the harder it will be to trust them when we've found them. UN Secretary General Antonio Guterres sees this as an "**existential risk to humanity."**

Generative AI applications like Chat GPT have made the technology the most hyped of our generation. However, researchers are skeptical it's anything more than fool's gold. Park '24 writes that

Park '24 [Hyun Jun Park. "The Rise of Generative Artificial Intelligence and the Threat of Fake News and Disinformation Online: Perspectives from Sexual Medicine." Investigative and Clinical Urology, vol. 65, no. 3, 1 Jan. 2024, pp. 199–199, https://doi.org/10.4111/icu.20240015. Accessed 28 Feb 2025]

With the advent of **generative artificial intelligence** (AI), the internet **has become a breeding ground for fake news and misinformation**. **The phenomenon** of **fake news and misinformation has had**

significant impacts across various sectors, including the world of **finance and politics**. A notable example occurred in mid-January 2023, when the spread of a false report stating that the SEC (U.S. Securities and Exchange Commission) had approved a spot-listed ETF (exchange-traded fund) caused volatility in Bitcoin prices [1]. In May 2023, an instance of generative AI being used to create a fictitious image of a building near the Pentagon in Washington D.C. engulfed in black flames, leading to turmoil in the U.S. stock market [2]. Additionally, fabricated images of a former U.S. president being arrested and a fashionably dressed Pope in a white puffer coat were examples of fake news created using AI-generated fake photographs [3,4].

In the early days of the pandemic, misinformation about the source of the virus and its spread was prevalent due to the scarcity of information about COVID-19. Fake news exploiting medical professionals' expertise has also damaged medical institutions and professional organizations.
The public is bombarded with information from frequently unreliable sources. Misinformation is a powerful destructive force in this age of global communication, when one false idea can spread instantly to many vulnerable ears [5].

The ease of replicating and editing everything from a person's face to their voice with just a few clicks signifies a new era where the boundary between reality and virtuality is rapidly dissolving. The increasing prevalence of AI-generated false information presents a significant challenge.

Generative AI mimics humans tricking students into falling for misinformation, and teachers over trust student's abilities letting them fall further into the trap.

Wineburg & Ziv '24 [Wineburg, Sam, and Nadav Ziv. "What Makes Students (and the Rest of Us) Fall for AI Misinformation?." Education Week, Editorial Projects in Education, 25 Oct. 2024, www.edweek.org/technology/opinion-what-makes-students-and-the-rest-of-us-fall-for-ai-misinformati o n/2024/10. Accessed 28 Feb. 2025.]

Four years ago during the 2020 election, we warned in the Los Angeles Times that **young people were struggling to spot disinformation because of outdated lessons on** navigating the internet. **Today, educators risk making the same mistakes with artificial intelligence. With the election at our doorstep, the stakes couldn't be higher.**

Previous work by our research team, the Digital Inquiry Group (formerly the Stanford History Education Group), showed that **young people are easily deceived because they judge online content by how it looks and sounds. That's an even bigger problem with AI, which makes information feel persuasive even when it fabricates content and ignores context.** Educators must show students the limits of AI and teach them the basic skills of internet search for fact-checking what they see.

When it comes to AI, leaders preach "great excitement and appropriate caution," as Washington state Superintendent Chris Reykdal put it in a recent teachers' guide. He writes of a "full embrace of AI" that will put that state's public education system "at the forefront of innovation." New York City schools former chancellor, David C. Banks, who stepped down amid a federal investigation, said in September that AI can "dramatically affect how we do school" for the better. The **"appropriate caution," however, remains a misty disclaimer.**

Washington state's guidelines, like California's, Oregon's, and North Carolina's, rightly warn that AI may be biased and inaccurate. Washington state stresses that students shouldn't automatically trust the responses of large language models and should "critically evaluate" responses for bias. But this is like urging students in driver's education to be cautious without teaching them that they need to signal and check blind spots before passing the car ahead of them.

This pattern repeats the mistakes we saw with instruction on spotting unreliable information online: **educators wrongly**

assuming that students can recognize danger and locate content that's reliable.

Massachusetts Institute of Technology professor Hal Abelson tells students that if they come across "something that sounds fishy," they should say, "Well, maybe it's not true." But students are in school precisely because they don't know a lot. They are in the least position to know if something sounds fishy.

Imagine a history student consulting an AI chatbot to probe the Battle of Lexington, as one of us recently tested. The large language model says this conflagration, which launched the American Revolution, was initiated "by an unknown British soldier." In truth, no one actually knows who fired first. The chatbot also reports that "two or three" British soldiers were killed during the skirmish. Wrong again. None was. Unless you're a history buff, this information doesn't sound "fishy."

A second danger is that **AI mimics the tone and cadence of human speech**, tapping into an aesthetic of authority. Presenting information with confidence is a trap, but an effective one: Our 2021 national study of 3,446 high school students reveals the extraordinary trust students place in information based on a website's superficial features.

When **students conflate style with substance and lack background knowledge**, the last thing they should do is try to figure out if something "sounds fishy." Instead, the detection of unreliable information and responsible use of AI rests on internet search skills that enable them to fact-check.

It is because we believe education should be on the frontlines of fighting misinformation, not introducing it to students, that we negate Resolved: In the United States, the benefits of the use of generative artificial intelligence in education outweigh the harms.

AI is being rushed into the classroom by large tech companies for profit- it is untested, produces faulty information, and is attempting to automate teachers out of jobs. Williamson '24 explains that

Williamson, Ben, et al. *Time for a Pause: Without Effective Public Oversight, AI in Schools Will Do More Harm than Good.* National Education Policy Center. March 2024, nepc.colorado.edu/sites/default/files/publications/PB%20Williamson_0.pdf.

In the spring of 2023, digital technology titans warned that artificial intelligence posed "profound risks to society and humanity,"[2] and called for regulation of its development and a pause in implementation of related applications.[3] But concurrently, **major players such as Google, Microsoft, Meta, and Amazon raced both to integrate AI into their platforms**[4] and **to fend off regulation**.[5] **Tech industry marketing went into high gear**[6] **and** soon the popular press overflowed with industry hype and speculation about AI's potential and pitfalls.[7] **The effect of all the** attention-grabbing predictions, self-interested corporate behavior, flamboyant **marketing** claims, and uncritical reporting **has been to obscure the** immediate **dangers posed by AI's rapid implementation**.[8] Given this backdrop, it is not surprising that rhetoric promoting **AI's alleged ability to positively transform teaching** and learning **has dominated** discussions about **its impacts on education**.[9] School administrators and teachers already use an array of digital educational technologies in teaching and management.[10] Their use has increasingly obscured educational decision-making, made a mockery of student privacy rights, and allowed student data to be exploited for non-school purposes.[11] **In the absence of** effective **public oversight,** the introduction of **AI systems** and applications **in education will** likely **intensify** these **problems** and create many more.[12,13] As existing school-focused platforms and applications are updated to include AI, the immediate danger facing educators is not a future apocalypse. Instead, the danger is that **AI** models and **applications will become enmeshed in school** processes and procedures **in ways that allow private entities to** increasingly **control the structure and content of public education**, to reinforce surveillance practices, **and to amplify existing biases and inequalities**.[14] For decades, academic researchers have worked on AI models for use in schools.[15] Today, however, it is **commercial enterprises** that **are aggressively pushing AI** (and its attendant risks) **into classrooms**.[16] The campaign to promote AI in education follows the logic of a half century of commercial, political, and ideological efforts to privatize and commercialize education.[17] Given this logic it is not surprising that, despite the known dangers, **corporations**, private researchers, **and government**s **are aggressively**

promoting the use of **AI**[18] **before a** statutory and **regulatory framework** has been put in place to ensure that AI programs are transparent and subject to effective public scrutiny and control.[19] **This puts schools under tremendous pressure to accept AI as** an **inevitable** upgrade to existing processes.[20] Computer scientists and software developers focus primarily on technical engineering questions[21] and corporate **leaders and investors prioritize profit**[22] over the common good. Nevertheless, **educators are being asked to trust** that these **people who have no educational expertise and** who stand to **financially benefit when AI is used in schools**, are best suited to imagine and lead educational transformation. III. Review of the Literature The term "artificial intelligence" (AI) was first introduced by computer scientists in the 1950s, though many of its underlying mathematical processes and mechanistic procedures can be traced back to early models of computing and manufacturing in the 1800s.[24] Standard definitions of AI usually refer to computers performing tasks that only humans could normally do.[25] Today, however, AI has become a slippery term without a widely agreed-upon meaning.[26] This enables marketers to apply the label "AI" to almost any digital process or product they are selling.[27] Development of Artificial Intelligence—From Rule-Based Systems to Generative AI AI has been developed over many decades by people and corporations with distinctive agendas and strategies.[28] Their efforts are poised to shape AI implementation in schools, especially since AI development and marketing efforts are rapidly expanding.[29] Three principal phases of AI's evolution began in the 20th century, when scientists and technicians focused on building systems that followed a sequence of pre-set rules, or algorithms, derived from the knowledge of experts.[30] Eventually, over the past two decades, a machine learning approach emerged. This approach involves using sophisticated mathematical processes (known as learning algorithms) to analyze massive amounts of data (known as big data) to identify trends or commonalities within it.[31] On the assumption that the vast amount of data analyzed confers validity on the trends identified, such programs generate predictions about behavior in complex events and phenomena—including human behavior.[32] The success of such analyses, however, has been mixed.[33] For example, one study asked six teams of data scientists to predict children's life outcomes with huge quantities of big data and cutting-edge machine learning tools. None of them came close to a reasonable level of accuracy, casting serious doubt on the use of AI in social policy areas like education.[34] Since late 2022, a further evolution of machine learning, generative AI,[35] has become an object of intense public, media, and political interest.[36] Generative AI deploys even more complex learning algorithms to create original text, images, and audio from data collected from the web or other sources.[37] This line of research has produced a series of program models often now called foundation models because they can be adapted for highly diverse purposes.[38]

Williamson continues that

Williamson, Ben, et al. *Time for a Pause: Without Effective Public Oversight, AI in Schools Will Do More Harm than Good.* National Education Policy Center. March 2024, nepc.colorado.edu/sites/default/files/publications/PB%20Williamson_0.pdf.

For example, **generative AI** large language models predict which word is most likely to follow the words preceding it by breaking down words into a sequence of numbers and then calculating the most probable response.[39] Large language models have provoked widespread excitement because their promoters have promised productivity gains for businesses as AI programs relieve workers of everyday writing tasks.[40] Similar support for teachers and students is also a promised benefit.[41] It is, however, **questionable that these promises will be realized in practice.**[42] For instance, while models often generate correct responses, computer programs **cannot understand what words mean**.[43] Therefore, the accuracy of their responses or the value of their responses in any given circumstance cannot be assumed.[44] Indeed, at the moment, **AI is often "stupid."**[45] **It invents facts mangles the results of analyses destabilizes information sources and produces dangerously wrong assertions about matters of social, public, and cultural importance**.[46] When Google demonstrated its Bard language application as an educational tool by getting it to answer questions about the James Webb telescope in 2023, for example, the model gave a false factual response, prompting a market plunge in Google's company value.[47] Concerns have also emerged about AI development processes. For technical and commercial reasons, AI models are not transparent—they are not publicly explained in any detail.[48] For example, an analysis of 10 leading foundation models found limited information about where data for analysis came from, how much computing was necessary to create the models, and what specifics were embedded in analytical algorithms.[49] Many machine learning models are black box models, meaning that their mechanisms are said to be too complicated to explain or not explainable at all.[50] Others are hidden from public view by proprietary rights accorded to corporations.[51] Black box foundational AI models are key to large technology corporations' plans to expand their proprietorial models

into all sectors, to grow global market share, and to generate maximum profit.[52] In the absence of effective public oversight and regulation, running generative AI programs is currently only possible by using "Big Tech" companies' databases, high-powered computing capabilities, and financial resources. This makes it likely that the proprietary AI models of a few corporations will become the foundation for the vast majority of AI applications developed.[53] Meanwhile, regulators and lawmakers will be left struggling to respond by creating a patchwork of after-the-fact regulatory protections.[54] AI in Education Since the 1960s, scientists and technology companies have explored ways to apply AI in education. AI in Education (AIED) is a major field of research and development.[55] The AI applications being promoted to schools today were preceded in the 1960s and 1970s by "Intelligent Tutoring Systems" and "Computer Assisted Instruction" systems.[56] Since the early 2000s, researchers have gathered, stored, and analyzed massive quantities of educational data with the intention of informing institutional and instructional strategies.[57] These approaches are now routinely considered synonymous with AIED, and have also been rapidly commercialized by the ed tech industry.[58] Most AIED applications employ big data and machine learning to produce various predictions and automated actions—such as predicting that a student may fail an assessment or creating a "personalized" intervention intended to produce a desired learning outcome.[59] Research on AI in education has developed and tested various approaches and reported modest effectiveness on measurable learning achievement—performance on quizzes and tests, for example.[60] Current excitement about its potential is motivating both public and private sources to generously fund researchers trying to find ways to improve learning outcomes using AI.[61] However, the assumption that AI in education can be understood primarily as a technical matter best addressed by scientists and companies is increasingly challenged by researchers who argue that a narrowly technical perspective may lead to both bad policy and bad pedagogy.[62] They point out that AI exists in social, economic, and political contexts that shape its development and uses.[63] How AI is adopted by different educational stakeholders (including AIED researchers, ed tech entrepreneurs, corporate leaders, and policymakers) will have significant implications for its use in schools.[64] **The fact** that entrepreneurs and **corporations** funded by venture capital and private equity **are rushing to promote AI in education will** inevitably **narrow possible applications** to those preferred by stakeholders with financial interests.[65] Small-scale ed tech start-ups and Big Tech corporations alike see AI as an opportunity,[66] leveraging popular hype to market such education products as personalized learning programs, automated lesson plan generators, and AI tutoring chatbots, called "tutorbots," to schools.[67] Compelling evidence for the effectiveness of tutorbots in education remains scarce,[68] though this does not prevent entrepreneurs and researchers from proclaiming their usefulness.[69] Policymakers routinely invoke AI rhetorically, calling on schools to embark on "digital transformation,"[70] often with little attention to social, economic, legal, or ethical implications.[71] These calls dovetail with existing political priorities on performance monitoring, account ability, efficiency, and effectiveness—all of which require extensive collection of data about students.[72] Although systems of test-based accountability have existed in schools since the 1990s,[73] they will expand and intensify as AI is used to

continuously monitor and assess student learning.74 As a result, commercial **AI systems will** increasingly serve as private actors in public education as schools, districts, and governments **relinquish key tasks**, functions, **and responsibilities** to third-party technology vendors.75 Existing and potential uses of AI in education are not merely innovative technical add-ons to teaching and learning practices or engineering solutions to schools' existing pedagogic and administrative problems. Rather, AI in education has been spurred by multiple forces: longstanding efforts by scientists to measure, predict, and support learning processes and outcomes; commercial aspirations to profit from selling products to schools; and the political objective of being perceived as having improved school efficiency and accountability as well as cutting costs. As things currently stand, these ambitions have begun to coalesce into a vision of AI-driven schooling in which commercial products assess student learning, automate teaching, and make decisions about student progress. Inadequate Research Base **Despite** the **extensive research in the field of AI in Education** (AIED) and the burgeoning research on machine learning, **there is** remarkably **little evidence to support** claims of **AI's ability** to "transform" schools.76 While AIED researchers have produced many research findings, their studies tend to focus primarily on measures of individual student engagement and performance (assessed by standardized achievements tests), or on "engineering" problems such as designing increasingly sophisticated algorithms and enhancing machine learning effectiveness.77 Overall, AIED studies tend to find ambiguous results, lack independence and scale, and fail to address more fundamental questions about educational goals.78 AIED research therefore often promotes a view of education transformation as improving measurable individual outcomes despite very limited evidence that AI "works."79 In effect, such studies reduce well-researched and nuanced theories of how humans learn to whatever can be made into a mathematical model (however complex), and they ignore the contested terrain of exactly which goals and curriculum public schools should embrace.80 Moreover, claims that AI can solve major educational problems—such as lack of qualified teachers, student underachievement, and educational inequalities—rely to a considerable extent on conjecture rather than evidence.81 Even more problematic are **the serious methodological flaws in machine learning research that call into question the validity of hundreds of studies.82 The nature of the flaws, in general, leads toward "over optimism" with respect to the usefulness and value of machine learning applications in a variety of fields**.83 These findings are particularly concerning because **they call into question** not only commercial marketing claims, but also t**he scientific evidence base supporting the widespread implementation of AI systems in** all sectors,84 including **education**. Finally, because of the very high computing costs associated with running machine learning models, most researchers have to rely on systems from the dominant AI companies themselves in order to conduct research85—the same corporations that often fund AI studies.86 This makes research dependent on corporate resources, funds, and business practices, giving AI firms considerable influence over not only AI development, but also the academic research that depends on their systems.87 It also compromises an important part of the research process, which is reproducing findings to verify their validity. When a company changes or stops supporting a particular model, researchers cannot reproduce studies conducted earlier.88 This renders the research base unstable and unverifiable—and thus unusable as a basis for assessing subsequent models. IV. Recent Developments In November 2022, OpenAI released ChatGPT, accelerating the race to develop and market generative AI platforms and applications. Companies that have developed artificial intelligence foundation models, including **OpenAI, Google, Meta, and Amazon**, **intend to expand rapidly** and "scale up" **in** every sector they enter—including **education**.89 While media attention has focused on students using ChatGPT, **these companies are swiftly expanding AI in education** in several ways: 1. Selling access to digital systems to schools. For example, Amazon sells schools and districts access to cloud computing facilities, enabling them to use its AI systems to analyze institutional and student data.90 2. **Adding AI features to products that schools already use**. Google, for instance, has begun introducing AI into its Workspace suite for schools, which includes the Classroom platform used by schools worldwide.91 3. **Integrating** AI applications into new or upgraded products. OpenAI, for example, partners with ed tech companies to integrate its language models into services it promotes as "**AI teaching assistants**,"92 and is also exploring its own educational applications of ChatGPT.93 4. Building AI into services such as search engines and other everyday applications commonly used in classrooms. For example, Microsoft introduced new AI applications in its Office software, based on OpenAI technologies, which are promoted for educational use.94 Current promotion of AI in education focuses largely on the pedagogical uses of AI applications that provide automated language and image producing capabilities.95 For students, reporting has, for example, focused on their use of automated tools, such as ChatGPT, to write assignments,96 **and** of **personalized learning "tutorbots" to mimic** a one-on-one tutoring experience.97 For teachers, it has focused on their use of AI "assistants" to create **lesson plans, develop grading systems, or review student progress**, among other tasks.98 Proponents assert that such AI applications offer students personal assistance in learning and offer teachers time-saving support.99 Entrepreneurial educators have produced guidebooks and training materials for teachers,100 and OpenAI has launched a Teaching with AI guide to train teachers to use its applications.101 Accepting proposed benefits as real, international and governmental organizations have supported the use of AI products in schools. For example, the Organisation for Economic Co-operation and Development (OECD), worried that AI will soon outperform humans on many cognitive tasks, has called for the urgent modification of formal education systems so students can learn skills to complement AI rather than skills for tasks that could soon be automated.102 The US Department of Education's Office of Educational Technology encourages teachers to involve themselves in developing and evaluating AI applications for education.103 **This** all **results in** teachers and **schools being pushed to accommodate untested** and opaque **commercial AI** applications **or risk being "left behind."** 104 **Recent regulatory proposals show little sign of slowing the rapid advance** of AI into **schools**.105 An executive order released by President Biden in October 2023 outlined directives to, among other things, increase federal oversight of foundation model testing, protect data privacy, and "promote innovation and competition."106 It mandates the Department of Education to create an "AI toolkit" incorporating "appropriate human review of AI decisions, designing AI systems to enhance trust and safety and align with privacy-related laws and regulations in the educational context, and developing education-specific guardrails."107 However, the fact sheet accompanying the executive order described the Department of Education's mandate differently as: "shap[ing] AI's potential to transform education by creating resources to support educators deploying AI-enabled educational tools, such as personalized tutoring in schools."108 Although the Department of Education will base its toolkit and guidance on its goal to "keep humans in the loop," two assumptions guide its approach: that "learning" is defined as and limited to those things that digital programs can measure and that AI programs can and should "optimize" learning as so defined. As noted earlier, despite the risks of rushing untested tech into classrooms and the lack of implementation criteria or regulatory controls,111 schools face unrelenting pressure to "modernize" by adopting artificial intelligence.112 While AI applications are marketed as ways to address teaching and learning problems and to streamline school administrative processes, they carry with them all the limitations, problems, and risks inherent in the AI models used to run them.113 **Rushing AI into schools increases the likelihood that these technologies will reproduce or intensify many problems, making any potential benefit less significant than the potential harms**. Two heavily promoted applications illustrate the perils of uncritical adoption of AI in schools: tutoring chatbots, based on large language models, which promise to personalize and automate teaching114; and adaptive learning platforms, which use large quantities of student data to make predictions, customize classroom resources and activities, and automatically intervene in pedagogical processes.115 The dangers that **such products** pose include costing teachers more time than they save, artificially restricting the definition of "learning," **sidelining teacher expertise**

and relationships with students, **introduc**ing **curricular misinformation** **and increas**ing**bias and discrimination in**classrooms and **schools**. In light of such dangers, before expanding or allowing AI applications into schools, policymakers and ducation leaders should consider what it would mean to use AI responsibly, whether its potential benefits outweigh the costs, and whether its adoption in education is truly inevitable. Dangers in Teaching and Learning Restricted Teaching and Learning Khan Academy has developed several offerings for technology-based schooling,116 expanding them beginning in March 2023 with heavy marketing of its Khanmigo "tutorbot" to parents, teachers, and school districts.117 Khanmigo's marketing claims that the tutorbot can engage in personalized and "conversational" pedagogic interaction with students, thus enhancing "personalized learning."118 Khanmigo incorporates OpenAI's GPT-4 language model to generate customized educational content, lesson plans, and assessment, and to perform such other tasks as monitoring student progress.119 Its marketing materials promise one-on-one "virtual tutoring" for students and a "personal teaching assistant" for teachers. These materials also claim that Khanmigo can understand and respond individually to students' academic and career goals and save teachers time by providing a content library for lesson planning, reports of student progress (along with "recommendations for what to do with that information"), and a "personal concierge service" that can help teachers "just do more."120 Such marketing suggests Khanmigo can make decisions that mimic a real teacher's thinking and responses, and so satisfy the needs both of students who require a teacher's attention and teachers who cannot offer that attention because of other demands on their time. It also reveals assumptions about the nature of teaching and learning embedded in the programming that are unavoidably imposed on real teachers, students, and communities that work with the product. Algorithms in personalized learning applications implement a narrow understanding of learning and a highly constrained pedagogic model.121 "Learning" in these applications typically consists of students engaging in computer-based activities and producing "correct" or acceptable answers within a limited range of predetermined responses.122 The applications present teachers with reports of student "progress" on the measurable tasks and suggest strategies to improve students' "performance."123

Any school adopting these applications tacitly accepts that learning can largely be reduced to a narrow range of behavioral responses and tacitly rejects more complex child-centered understandings of learning.124 The theory of learning built into AI chatbot applications like Khanmigo prioritizes data-based numerical profiles of students—considered "objective" measures of performance—while undermining teachers' ability to make professional judgments about their students and devaluing teachers' subjective experience, subject matter expertise, classroom interactions, and contextual knowledge of a given child, class, or social setting.125 A supposed advantage in such AI-based programs is that they reduce administrative burdens on teachers and allow them more time to teach. Khanmigo's advertising promises, for example, "your evenings are yours again without compromising quality."126 This is persuasive marketing for**overburdened teachers**, who might**welcome outsourcing** their overwhelming administrative duties to a computer program, or might be relieved to occupy students with technology that theoretically supports their learning while they wait for teacher time in an overcrowded classroom. Delegating tasks to a tutorbot,**however**, inserts a digital intermediary between students and teachers who are pedagogical experts in their fields and who know their students and understand their context. Moreover, despite **generative AI**'s humanlike communication, it **is limited**to responding to prompts and queries that fall within a product's established parameters. As a result,**teachers must devote time to understand**ing and mastering such **limitations and** then **teach students how to ask questions and verify responses**.127 **Even as AI automates some** administrative **tasks**, then, **it is likely to introduce other time-consuming pressures and burdens on teachers.**128 Rather than reducing demands, A**I of the kind currently promoted by private commercial enterprises can add further complexity to teachers' workloads by ultimately positioning them as servants of the technology tasked with ensuring its smooth classroom operation**. Reconfiguring the core systems of education—instruction, curriculum, and assessment—to accommodate AI will demand laborious efforts by educators to adjust their professional practices. At the same time, it will deny them a voice in determining whether proposed changes have real value and should or should not be implemented. Academic AIED research and development, informed by partnership with teachers and students, will struggle to be heard above the current promotional and speculative rhetoric around commercial AI applications for schools.129 AI therefore presents a critical challenge to the institutional autonomy of schools' decision-making processes, and particularly to teachers' autonomy to make professional decisions about their pedagogic practices.Curricular Misinformation A key challenge of AI for teachers is that **generative AI technologies** are trained to**produce text that seems convincing even though**it might contain **false** information.130 Applications that use **generative AI** to support teachers' lesson planning and resource generation, for example, could **flood the classroom with** misleading inaccuracies or **false information**.131 **This is not a problem likely to be resolved**, because **the data such programs depend on** may**degrade further as automated content spreads across the web.**132 The danger is that the information environment will be overrun by AI-generated text, making it impossible to ascertain the authority or authenticity of any online source,133 and therefore **rendering online sources useless** or misleading **for education**al purposes.134 As has been documented in both language- and image-generating AI,135 instabilities and errors in underlying data make AI chaotic and unreliable.136 No doubt this is why Khanmigo recommends that users not rely on its responses but rather verify its information using such resources as "textbooks, articles, or other trusted sources."137 And why it also restricts the time students can spend interacting with the tutorbot, noting that "extended interactions are more likely to lead to poor AI behavior."138 These recognized limitations beg the question: Why spend time using this application or any AI application when it seems obvious their use will result in more work and very likely undermine effective teaching?139 Outstanding questions remain, then, about ed tech and Big Tech market leaders becoming powerful gateways to online learning content. Do their AI applications improve the quality of information taught in the classroom? How much should the content produced by AI, having been trained on material scraped from the web by companies like OpenAI, be trusted by teachers and students? What are the implications of students encountering "poor AI behavior" from a tutorbot? It is also not clear whose responsibility it must be to check the quality of the content produced by AI. Khanmigo implicitly makes already overworked teachers responsible for doing the hidden labor of checking for errors and providing feedback that the product's programmers can then fix one-by-one.140 This scenario makes it very likely that such errors will be overlooked and that the accuracy of information that students encounter via AI will degrade over time.141 AI poses a real threat to the accuracy of school knowledge and thus to the validity of curriculum materials. In a context where students and many teachers already source content online (for example, through sites such as Teachers Pay Teachers142) it may become increasingly difficult to tell an authoritative and accurate source from a plausible but fallacious one produced by AI. This is likely to become an even bigger problem with OpenAI's release of "GPTs"—versions of ChatGPT that can be customized by users—as individuals are able to create their own educational AI applications and share them freely online, with few quality checks and controls.143 Potentially Amplified Bias and Discrimination Because **AI models** are trained on either internet data or historical data, they **incorporate biases that** can **transfer to their educational applications.**144 For example, when ChatGPT was released in November 2022, educators expressed anxiety about students cheating on written assignments, prompting ed tech companies to develop automated AI detectors. The makers of**Turnitin**, a product already used internationally to detect student plagiarism, added**AI detection**functionality in early 2023, arguing it would be able to detect distinctive markers of AI-generated text.145 However, independent studies have found that such AI**detectors are prone to error**, **leading to a surge in false accusations of cheating**.146 Such accusations are

disproportionately**biased against non-native English speakers** who tend to **write in simpler sentences** that **AI flags as suspicious.** 147 Similarly, the **automatic AI** essay **grading programs used in many states are prone to bias against certain demographic and ethnic groups**, falsely awarding high grades to work that features sophisticated language and structure, regardless of the meaning or quality of the writing.148 **This is because AI programs do not "know,"** in any real sense, **what makes for a good or bad essay**, but can only search for patterns that correlate with higher or lower human-assigned grades. Likewise, GoGuardian, which uses algorithms and AI to monitor students' social media and web browsing activity for "suspicious" content or behavior, routinely flags multiple categories of non-explicit material as harmful or dangerous—including material on general educational sites.

Additionally, teachers are also getting misinformed due to GAI

Maddy **Dwyer**, Elizabeth Laird, 3-27-**2024**, "Report – Up in the Air: Educators Juggling the Potential of Generative AI with Detection, Discipline, and Distrust", Center for Democracy and Technology, https://cdt.org/insights/report-up-in-the-air-educators-juggling-the-potential-of-generative-ai-with-detection-discipline-and-distrust/ / PS

"**Educators are having a very different experience with generative artificial intelligence (AI) since the 2022-23 school year came to a close.** K-12 schools have now had the opportunity to take a breath and regroup to determine how to get a grip on the explosion of generative AI in the classroom – after the education sector was caught off guard when ChatGPT burst abruptly onto the scene during the last school year.

To understand how teachers are currently interacting with and receiving support on this technology, the Center for Democracy & Technology (CDT) conducted a nationally representative survey of middle and high school teachers in November and December 2023. This research builds on previous CDT findings that highlighted how schools were failing to enact and/or share policies and procedures on **generative AI** and how, as a **result**, **teachers lacked clarity and guidance**, were more distrustful of students, and reported that students were getting in trouble **due to this technology.**"

The impact is critical thinking

Rushing faulty generative AI into the classroom will irreparably damage students' critical thinking. Basha '24 explains that

Yunus Basha.J. The Negative Impacts of AI Tools on Students in Academic and Real-Life Performance.Int.J.Soci.Sci.Vol.1(3).2024.Pp:1-16.
https://doi.org/10.51470/IJSSC.2024.01.03.01.
https://nsfjournals.com/article-files/pdf/WyWmemNe8hyG6vgiQoPO-1723214664.pdf

As the usage of **AI** tools escalate **in the educational setting**, it's crucial to address the negative impacts on students academic performance. The Students activities and involvements in academic domain plays a vital role for their holistic development, the adoption of AI tools in the academic activities can be encourage until it will become more reliant. Over-reliance on AI tools by the students for the achievement of their academic goals **can** be result **decline** in **critical thinking skills**, disappearance of the traditional skills such as handwriting skills, **memorization and math** solving **skills** which are still important for academic settings. Moreover, the over-reliance of AI tools may lead to concern about ethical issues such as privacy of the students these AI tools may collect the data of the students without their consent and knowledge which can be possible to misuse their sensitive information, therefore addressing ethical concerns should not be ignore in the view of to establish safeguards to students well-being and rights. Furthermore, **AI tools** may **contain inaccuracy and biases** in the readily available data **which** can **affect the academic work of the students** that may lead to **get** wrong conclusion and **false information from AI powered platforms**. AI tools provides vast amount of information to **students**

which can leads to the students use to **copy and paste the data without understanding the content**s or the topic,**this** attitude of the students makes them lethargic and inactiveness moreover it will lead to hindrance for cognitive development of the students which is vital for academic Settings. AI powered applications or tools are specifically designed to enhance learning experiences

Indeed, rather than using AI to supplement learning, students over-rely on it. Kelly and Smith '24 find

Maj. Patrick **Kelly and**, U.S. Army Maj. Hannah **Smith**, May 20**24** Online Exclusive Article U.S. ArmyArmy University Press, "How to Think About Integrating Generative AI in Professional Military Education," https://www.armyupress.army.mil/Journals/Military-Review/Online-Exclusive/2024-OLE/Integrating-Generative-AI/, accessed 2-13-2025 // noah However, the problem with universal adoption is that **generative AI** can do much more. ChatGPT **can synthesize evidence**, **generate** a **thesis**statement, **and develop a** coherent and **comprehensive argument**to support it, all while balancing contradictory information. Within the context of CGSC, a student might upload doctrine files and ask ChatGPT to produce a summary of selected excerpts. And in these capacities, ChatGPT is clearly more than the latest typewriter. The differences are apparent when we attend to specific use cases and remind ourselves that knowledge is not only a product but also a process.10 Sometimes, **ChatGPT** can complement critical thinking by providing information and efficiencies. But other times, the technology **can substitute for critical thinking**.**Skeptics dismiss**this point too quickly based on **ChatGPT's known** limitations, like tendencies **to "confidently" misinform, "hallucinate" facts, or fail to cite sources. However, the key point in** this debate **is** not whether generative AI can perform all written tasks better than humans, or whether a chatbot is indeed intelligent, **but that it can consistently** **produce passable work**, **which** creates incentives forstudents to **shortcut** processe**s** of **discovery and deep thinking**.

It's been proven. Knapp '25 reports that

Knapp, Alex. "The Prototype: Study Suggests AI Tools Decrease Critical Thinking Skills." Forbes, 10 Jan. 2025, www.forbes.com/sites/alexknapp/2025/01/10/the-prototype-study-suggests-ai-tools-decrease-critical-t hinking-skills/. Accessed 5 Mar. 2025.

New AI tools are slowly becoming ubiquitous, being added to the software and hardware we use every day (sometimes whether we like it or not). But if we're using artificial intelligence to perform tasks, search for information and solve problems, what does that mean for the intelligence we're born with? To figure this out, a team of researchers conducted **a study involving 666 individuals** ages 17 and up, representing a diverse population. It first evaluated the extent to which each of them made use of AI tools, then tested their critical thinking skills. The results of the study, which were published in the journal Societies, **found** that **those who used AI tools a lot showed worse critical thinking** abilities than those who didn't use them often or at all. **Whether someone used AI** tools **was a bigger predictor of a person's thinking skills than any other factor, including educational attainment. The reason** for this **is** a phenomenon called "cognitive **offloading**" – where **people's thinking and problem-solving are** essentially **delegated**. Frequent cognitive offloading reduces a person's ability to independently think and solve problems. "**This relationship underscores the dual-edged nature of AI technology**," the study authors wrote. "While it enhances efficiency and convenience, it inadvertently fosters dependence, which can compromise critical thinking skills over time."

Additionally, in a peer reviewed study,

Fonkam et al. 24 [Mathias Fonkam, PhD in computer science and Associate Teaching Professor @ Penn State University with over 20 years of experience in computer science education, xx-xx-2024, Risks of AI-Assisted Learning on Student Critical Thinking: A Case Study of Albania, International Journal of Risk and Contingency Management, https://www.igi-global.com/article/risks-of-ai-assisted-learning-on-student-critical-thinking/350185] BZ

The analysis of the data also revealed several key findings regarding the impact of reliance on AI tools for assignments on students' problem-solving skills. The descriptive statistics as shown in Table 8 demonstrate that the mean reliance on AI tools for assignments was 2.17, with a standard deviation of 1.25. In comparison, the mean score for problem-solving skills was 3.06, with a standard deviation of 1.39. These statistics indicate a moderate level of reliance on AI tools and a slightly above-average self-assessment of problem-solving abilities among students. The Pearson correlation analysis revealed a correlation coefficient of -0.712 between reliance on AI tools and problem-solving skills, with a p-value of less than 0.000000001. This indicates a statistically significant negative relationship, suggesting that as students rely more on AI tools for their assignments, their problem-solving skills tend to decrease.

Further examination through linear regression analysis showed that the model explained 50.7% of the variance in problem-solving skills, with an R-squared value of 0.507. The regression coefficient for reliance on AI tools was -0.7918, with a p-value of less than 0.001. This coefficient indicates that for each unit increase in reliance on AI tools, there is an associated decrease of approximately 0.7918 units in problem-solving skills. The significant negative coefficient supports the hypothesis that increased reliance on AI tools negatively impacts problem-solving skills. The scatter plot shown in Figure 3 shows the correlation between reliance on AI tools for assignments and problem-solving skills. The scatter plot shows individual data points, and the black regression line indicates the negative correlation between the two variables. As reliance on AI tools increases, problem-solving skills tend to decrease, supporting the hypothesis (H3). Hypothesis testing using a t-test compared the levels of problem-solving skills among different levels of reliance on AI tools. The t-statistic was -2.618, with a p-value of 0.011, which is below the conventional threshold of 0.05. This result allows us to reject the null hypothesis and accept the alternative hypothesis (H3), confirming that reliance on AI tools for assignments significantly negatively impacts students' problem-solving abilities. The significant negative correlation and regression results underscore the importance of addressing the balance between using AI tools for efficiency and maintaining the development of independent problem-solving capabilities. These insights are critical for educators and policymakers aiming to integrate AI tools into educational practices without compromising essential cognitive skills.

This shows that for the 74 million students across the country, GAI is a threat to their education and future.

C2 Academic Dishonesty

AI substantially increases cheating - students who don't cheat also feel forced to do the same.

Akbari, N. Noor Akbari is the cofounder of rosalyn.ai, an exam proctoring service. (20**24**, October 24). The AI cheating crisis: Education needs its Anti-Doping Movement (Opinion). Education Week. Retrieved February 8, 2025 from https://www.edweek.org/technology/opinion-the-ai-cheating-crisis-education-needs-its-anti-doping-movement/2024/02//SJLH

The rise of "AI doping" is strikingly similar. **In a Study.com survey of 1,000 college-age students, 89 percent of respondents admitted to using ChatGPT to complete a homework assignment. Another 48 percent admitted to using it on at-home tests or quizzes, and 53 percent had the bot write an essay.** Yet, 72 percent of the students reported believing ChatGPT should be banned from campus networks.

Are they hypocrites? No. When enough players in a competitive game can cheat with a high upside and low risk of consequences, other players will feel forced to cheat as well. As Lance Armstrong told the French newspaper Le Monde in 2013, several months after publicly admitting he had used performance-enhancing drugs, it was "impossible to win the Tour de France without doping."

If enough students improve their grades using ChatGPT, their peers may conclude that it's "impossible" to compete unless they cheat, too. In cycling, not doping meant losing competition winnings and sponsorships. **In education, not using generative AI could mean losing out on college admissions, scholarships, and career opportunities.**

At the height of the sports doping crisis, a common argument was to let doping happen. If no substance is banned, isn't the playing field level? The counterargument is that substances are banned because they pose a health risk to athletes. If sporting organizations not only allowed but tacitly encouraged athletes to dope, the resulting biochemical arms race would have a sure loser: athletes and their well-being.

Likewise, the normalization of AI doping would create an arms race among students, resulting in several consequences for them and society:

1. Unchecked use of AI renders education pointless. We fund public education as a common good because it empowers citizens to live fulfilling lives and contribute to their communities. If the point of education is merely to get a diploma, then who cares if a student or AI does the work? The true point of an education, however, is to train a person's mind and character.

Claiming that students no longer need to learn skills like writing—because AI does it—is like arguing that no one should strength train because carts and forklifts move heavy stuff for us. This conflates means and ends. People lift weights for the inherent benefits to their mind and body. Likewise, we learn to write for the inherent benefits to our cognition and communication skills.

2. AI threatens to undermine academic integrity, the foundation for professional credibility. We trust our surgeon, certified public accountant, or lawyer because we trust the institutions that test, certify, and employ them.

Well, researchers have found that ChatGPT can pass the U.S. medical licensing exam, CPA exam, and bar exam. You might argue that cheating on those exams is almost impossible. But what about an online nursing exam, an online certification in cybersecurity, or an online degree in social work? A person who cheats for the credential in those cases could become a danger to others.

3. The struggle to maintain academic credibility could produce a two-tiered education system that is even more inequitable than the current one. Elite colleges with full-time professors and graduate students have the resources to design assignments in which AI provides no edge. Community colleges and online education platforms don't have that luxury. A class with hundreds of students and one part-time instructor cannot convert every digital test into an original research project or in-person test with pencils and paper—not without raising costs considerably.

Efforts to democratize education will be laughable if the only credible degrees come from private, in-person institutions that cost students an average of almost $56,000 per year.

K-12 schools will experience similar inequities, particularly between expensive private schools and crowded public schools. Schools with lower student-teacher ratios are better positioned to design assignments that limit students' reliance on AI than schools with more limited staff and resources.

So how do we address AI's threat to academic integrity and an affordable education?

Forget watermarking AI-generated text and AI detectors—they're easily duped. And forget academic "honor codes." In my birth country of Afghanistan, the strict honor codes of Islam that forbid corruption didn't stop the country from becoming an epicenter of corruption after the United States' 2001 occupation injected billions of dollars. Like Lance Armstrong, no one struggled to justify corrupt behavior when everyone else was doing it, too.

Exams, whether in person or online, must be proctored such that no one can cheat using AI. That said, to prepare students for the working world, schools should teach generative AI in classrooms using versions with limited capabilities.

The silver lining of the 1998 Tour de France was that the International Olympic Committee formed the World Anti-Doping Agency the following year. Though far from perfect, WADA created a unified list of banned substances and standards for detecting them. In other words, the organization defined what "doping" means in sports. Soon enough, education systems may need a WADA-like organization to define cheating in the AI age and set standards for preventing and detecting it.

Doping in sports undermined the fairness and meritocracy of a beloved institution, until that institution took the threat seriously. It's time we take AI doping in schools seriously.

Specifically, teachers are finding it difficult to detect the difference between human and AI output. Geist '24 confirms:

Geist, A. (2024). Readers can't accurately distinguish between AI and human essays, researchers find. [online] Yale Daily News. Available at: https://yaledailynews.com/blog/2024/11/18/readers-cant-accurately-distinguish-between-ai-and-human-essays-researchers-find/ [Accessed 23 Feb. 2025]. // JA CCHS

Research suggests readers struggle to tell the difference between human and artificial intelligence-generated essays. In a project organized by four researchers, including three from the School of Medicine, researchers tasked readers with blindly reviewing 34 essays, 22 of which were human-written and 12 which were generated by artificial intelligence. Typically, they rated the composition and structure of the AI-generated essays higher. However, if they believed an essay was AI-generated, they were less likely to rank it as one of the overall best essays. **Ultimately, the readers only accurately distinguished between AI and human essays 50 percent of the time,** raising questions about the role of AI in academia and education. "How would we even know, other than the word of the author, whether the paper was assisted by generative AI?" Dr. Lee Schwamm, associate dean of digital strategy and transformation at the School of Medicine, said.

And, there is no tool that can perfectly detect AI as Shabanov '24 highlights:

Ilya Shabanov (2024). How reliable are AI detectors for academic text and should you use AI for writing? - The Effortless Academic. [online] The Effortless Academic. Available at: https://effortlessacademic.com/how-reliable-are-ai-detectors/ [Accessed 23 Feb. 2025]. // JA CCHS

On average, numerous studies support the claim **that AI detection is unreliable.** Here is the accuracy of various tools from one of the most cited AI detection studies (Weber-Wulff 2023): The authors conclude their work with a stark warning that academics shouldn't rely on any of these tools: Our findings do not confirm the claims presented by the systems. **They too often present false positives and false negatives.** Moreover, it is too easy to game the systems by using paraphrasing tools or machine translation. Therefore, our conclusion is that the systems we tested should not be used in academic settings.

Even IF it doesn't get caught by teachers or AI tools, Cheating ruins academic performance in itself.

Sparks 11 [Sarah D Sparks, reporter and data journalist for Education Week, 3-29-2011, "Studies Shed Light on How Cheating Impedes Learning", Education Week, https://archive.ph/NT7jn, accessed 3-3-2025] RNM

Emerging evidence suggests **students who cheat** on a test **are more likely to deceive themselves into thinking they earned a high grade on their own merits, setting** themselves **up** for **future academic failure.**

In four experiments detailed in the March *Proceedings of the National Academy of Sciences*, researchers from the Harvard Business School and Duke University found that **cheaters pay for the short-term benefits of higher scores with inflated expectations for future performance.**

The findings come as surveys and studies show a majority of students cheat—whether through cribbing homework, plagiarizing essays from the Internet, or texting test answers to a friend's cellphone—even though overwhelming majorities consider it wrong. The Los Angeles-based Josephson Institute Center for Youth Ethics, which has been tracking student character and academic honesty, has found that while the number of students engaging in specific behaviors has risen and fallen over the years, the number of students who admit to cheating on a test the previous year has not dipped below a majority since the first study in 1992. In the most recent survey, conducted in 2010, the study found that a majority of students cheat sometime during high school, and the likelihood of cheating increases the older students get.

Of a nationally representative sample of more than 40,000 public and private high school students responding to the survey, 59.4 percent admitted to having cheated on a test—including 55 percent of honors students.

In addition, more than 80 percent of the respondents said they had copied homework, more than one-third had plagiarized an Internet document for a class assignment, and 61 percent reported having lied to a teacher about "something important" at least once in the past year. By contrast, only about 20 percent of students surveyed reported having cheated in sports.

"One of the sad phenomena is that, on average, one of the things they are learning in school is how to cheat," John Fremer, the president of consulting services at Caveon LLC, a private test-security company in Midvale, Utah, said of students.

While most academic interest in cheating has focused on how students cheat and how to stop them, the Harvard-Duke study adds to emerging research suggesting that the mental hoops that **students must leap through to justify or distance themselves from cheating can cause long-term damage to their professional and academic habits**. The findings also point to aspects of school climate and instructional approach that can help break the cycle of cheating and self-deception.

"We see that **the effect of cheating is, the more we engage in dishonest acts, the more we develop these cognitive distortions—**ways in which we neutralize the act and almost forget how much we are doing it," said Jason M. Stephens, an assistant professor of educational psychology at the University of Connecticut, in Storrs, who studies cheating among secondary school students.

Moreover, the more **students learn to focus on grades for their own sake, rather than as a representation of what they have learned**, the more comfortable they are with cheating.

Mr. Stephens, who was not involved in the Harvard-Duke study, quoted one high school student, "Jane," who insisted that cheating on a test does nothing to lessen the value of the grade. "It says an A on the paper and you don't go, 'Oh, but I cheated.' You're just kind of like, 'Hey, I got that A,' " she said.

That, said Zoë Chance, the lead author of the Harvard-Duke study, is where cheaters start lying to themselves.

Furthermore, this has translated into an <u>increase in plagiarism</u> by students. Chen et al '24 explain:

Chen, B., Lewis, C.M., West, M. and Zilles, C. (2024). Plagiarism in the Age of Generative AI: Cheating Method Change and Learning Loss in an Intro to CS Course. Plagiarism in the Age of Generative AI: Cheating Method Change and Learning Loss in an Intro to CS Course. doi:https://doi.org/10.1145/3657604.3662046.   // JA CCHS
Background: ChatGPT became widespread in early 2023 and enabled the broader public to use powerful generative AI, creating a new means for students to complete course assessments. Purpose: In this paper, we explored the degree to which generative AI impacted the frequency and nature of cheating in a large introductory programming course. We also estimate the learning impact of students choosing to submit plagiarized work rather than their own work. Methods: We identified a collection of markers that we believe are indicative of plagiarism in this course. We compare the estimated prevalence of cheating in the semesters before and during which ChatGPT became widely available. **We use linear regression to estimate the impact of students' patterns of cheating on their final exam performance. Findings: The patterns associated with these plagiarism markers suggest that the quantity of plagiarism increased with the advent of generative AI, and we see evidence of a shift from online plagiarism hubs (e.g., Chegg, CourseHero) to ChatGPT. In addition,** we observe statistically significant learning losses proportional to the amount of presumed plagiarism, but there is no statistical difference on the proportionality between semesters. Implications: **Our findings suggest that unproctored exams [have] become increasingly insecure and care needs to be taken to ensure the validity of summative assessments.** More importantly, our results suggest that generative AI can be detrimental to students' learning. It seems necessary for educators to reduce the benefit of students using generative AI for counterproductive purposes.

This is detrimental for two reasons — the first is trust between students and teachers. Talbert '06 confirms:

Talbert, R. (2006). Four reasons why academic dishonesty is bad. [online] The Chronicle of Higher Education. Available at: https://www.chronicle.com/blognetwork/castingoutnines/four-reasons-why-academic-dishonesty-is-bad [Accessed 23 Feb. 2025]. // JA CCHS
Academic dishonesty makes student assessment unreliable. When I give a piece of work to be graded, the reason I do so is that the grades I get back constitute data -- data from which can extract information about how students are doing and adujst my teaching accordingly. If a student fakes his or her own work, through plagiarism or cheat notes or whatever, the data that I get don't tell me truthful information that I can act upon. **More importantly, the intellectual needs of students go unmet because according to the data, everything is going well, when everything is not going well.** One corollary of this idea: It is not the case that academic dishonesty hurts only the students involved. It hurts all the students in the class who have the same needs as the cheater. **Academic dishonesty erodes the mutual trust between students and faculty that is at the core of higher ed. Students trust faculty to be knowledgeable in their fields and truthful in their teaching.** Faculty trust students to give true information about their progress via graded work. In other words, I'll teach you what you truly need to know, and you tell me truly how you're doing in your work. That combination of truthfulness and trust should create a sort of spiral where one by one, student misunderstandings are removed and true mastery is attained. **A breakdown on either end corrupts the entire process. And it only takes one instance of a breach of trust. If a student cheats or plagiarizes, that student's work is suspect for the duration -- even if the student learned his/her lesson and the suspicion is misplaced.**

Trust is an essential part of learning, and losing it ruins the foundation of education for millions of students. Varthana '23 confirms:

Varthana. (2023). Why is Building Trust Between Teachers and Students Essential? [online] Available at: https://varthana.com/school/why-is-building-trust-between-teachers-and-students-essential/ [Accessed 23 Feb. 2025]. // JA CCHS
Trust is fundamental in a teacher-student relationship because it forms the basic foundation of learning. When students believe in their teacher, they feel safe to ask questions, share their thoughts, and take risks without judgement of others. Trust [because it] fosters open verbal exchange, allowing teachers to provide constructive feedback and support each student's needs. It additionally encourages collaboration and mutual appreciation, growing a sense of belonging within the classroom. **With trust, students are more inspired to actively engage in lessons and participate in class discussions.** Moreover, it strengthens the bond among teacher and student, leading to a supportive and nurturing learning environment where students can thrive each academically and emotionally. Through the development of trust and nurturing a healthy relationship, teachers and students begin to value each other's individuality, show mutual respect, and exhibit polite behavior. Positive relationships with students help create a safe learning environment within the classroom. **In terms of academic learning, trust significantly promote students' acquisition of knowledge. Trust in the academic educational process goes beyond gaining**

Stephen L. Chew, Phd, 8-31-2023, "Student trust in the teacher: A critical but overlooked factor in student success", https://apa.org, https://www.apa.org/ed/precollege/psychology-teacher-network/introductory-psychology/student-trust //PS

"The policies, though within the purview of the instructor and applied uniformly, undermined student effort and learning. The most obvious issue **with** these policies is that they pitted doing the right thing for the health of the students against doing well in the course. There is, however, a deeper issue. The policies **undermined** student **trust** in the **teacher**. They **assumed a negative view of students, that they are unmotivated**, eager to skip classes, and seek easy ways to pass. Second, the policies also **prioritized compliance without regard to** the life stresses experienced by **students**. The policies were both **demeaning to students and dismissive of the demands in their personal lives.**"

The second reason is habit formation and job security as Meazure '23 concludes:

Meazure Learning (2023). Meazure Learning. [online] Meazure Learning. Available at: https://www.meazurelearning.com/resources/what-are-the-negative-impacts-of-cheating-on-students-in-college#:~:text=Impacts%20on%20Students%20Who%20Cheat,to%20their%20degree%20or%20program. [Accessed 23 Feb. 2025]. // JA CCHS

The same Turnitin study referenced previously asked students and faculty to determine whether certain behaviors counted as cheating. The study found that 18% of students did not think the use of unauthorized materials on an exam counted as cheating. A staggering 31% of students believed that hiring someone to write an essay for them is not cheating or is only somewhat cheating. Whether or not a student views a behavior as cheating can have direct impact on how likely they are to engage in behavior that may lead to a culture of cheating at an institution. How

Test-Takers Can Be Affected by Cheating on Exams The impacts of cheating on test-takers can be varied and consequential, both on those who engage in cheating behavior and those who don't. For those who do cheat, the most obvious consequences are academic. Cheating on an exam can result in failing that test or—in many cases—failing the class entirely. Academic consequences can go beyond one class as well, sometimes leading to suspension or expulsion from the program or the institution entirely. **Impacts on Students Who Cheat If allowed to continue,** cheating behaviors can be habit-forming [and]. **Even if a test-taker feels that cheating is wrong, they may be tempted to engage in cheating behavior on a specific exam or assignment that is of significant value to their degree or program.** If they give in to that temptation, it will be easier to do so the next time. **Like many forms of dishonesty, cheating** can be a slippery slope that may lead to more substantial behaviors in the long run. When test-takers who cheat on exams are not found out, the effects can spread out into the wider world, beyond the test, class, and college degree. **When a test-taker who cheats passes their exam, there can be dire consequences for their employment. They may find themselves hired for a job they are unqualified for, which could lead to a variety of unwanted results. Being unqualified could cost them a promotion, an assignment, or even their job.** From the employer's perspective, hiring an unqualified candidate leads to increased acquisition and firing costs, requiring them to replace incompetent workers. It could also lead to legal and financial repercussions for both employee and employer in a high-stakes job in an industry like medicine or finance, ultimately eroding the employer's confidence in the exam program, "The impacts of cheating on a program's credibility and the students in that program can be devastating, but a robust academic honesty policy and clear communication of expectations can help prevent a culture of cheating. " Ashley Norris, PhD, SVP, Strategic Communications and Policy According to a recent study conducted by the ATP Workforce Skills Credentialing Security and Privacy Committee, over 80% of employers surveyed agreed that it is very important to be able to verify credential authenticity. If there is a history of test-takers cheating on a particular exam or a particular program's exams, employers are much less likely to trust in that program's credentials, which ends up hurting everyone associated with that program—test-takers and instructors included. Impacts on Students Who Don't Cheat The effects of cheating don't only touch those who engage in cheating behavior. The reputational damage done to a program through a culture of cheating affects everyone in the program. In addition, students who don't cheat can face unique consequences when cheating is allowed or perceived to be allowed. When faculty believe there is a culture of cheating, they adjust their policies and responses accordingly. Often instructors will begin to police student behavior much more strictly when there is a perceived culture of cheating. This can mean putting undue pressure on academically honest students, causing them unnecessary stress and anxiety. This can negatively affect their performance on the exam and in the program as a whole. When students believe there is a culture of cheating, it can affect the way they view their program and their participation in it. With enough of a change in this view, students may leave the program entirely. Students who believe their peers regularly cheat may begin to feel they are at a disadvantage if they do not also cheat. It can also lead them to question the accuracy of their own assessments. They may begin to feel that in order to succeed in their program, they also need to cheat. This can lead to the moral disengagement we discussed previously.

Thus we negate.