The Future of Expertise in the Age of AI

Reflections on Stanford University's Tutor Co-pilot

Dr Neil Hopkin
Director of Education
Fortes Education

The Cambridge Consultancy Group
Leading Education Series 2024

The Rise of AI in Transforming Expertise Delivery

Stanford University has very recently announced Tutor CoPilot, a novel Human-AI approach that leverages a model of expert thinking to provide expert-like guidance to tutors as they tutor. Their study (Wang et al, 2024) presents the first randomised controlled trial of a Human-AI system in live tutoring, involving 900 tutors and 1,800 K-12 students from historically under-served communities.

Picture a young tutor, Emily, nervously stepping into her first classroom. Her heart races as she glances at the expectant faces of her students, their curiosity mixed with a hint of skepticism. But Emily isn't alone. Just as she pauses, unsure of how to explain a tricky concept, a small prompt appears on her screen—a suggestion from an AI guide known as the Tutor CoPilot. "Try using a story to illustrate the point," it advises. Suddenly, Emily feels a surge of confidence. She follows the suggestion, weaving a narrative that captures her students' attention. A small, almost imperceptible smile forms on her lips. This was just one moment, but it encapsulates a larger question: Can AI truly transform how expertise is scaled and shared in the most human of professions?

This scenario, while novel, is becoming increasingly common as AI tools like Tutor CoPilot attempt to redefine the relationship between human expertise and machine assistance. The promise is tantalizing: to democratize access to high-quality instruction by amplifying the capabilities of even novice educators. Yet, as Daniel Kahneman once observed, human decision-making is not purely rational; it is nuanced, deeply rooted in experience, and prone to biases that shape how we react under pressure (Kahneman, 2011). In education, these biases can be both a flaw and a strength. AI, with its data-driven precision, can offer a counterbalance—flagging potential pitfalls that a novice might miss. But can it capture the subtle, almost invisible interplay between intuition and knowledge that experienced educators wield so effortlessly?

The challenge of replicating this expertise is at the heart of what Howard Gardner described in his theory of multiple intelligences: human intelligence is not a monolith but a collection of varied, interdependent abilities (Gardner, 1983). A teacher might rely on logical reasoning to solve a math problem but switch seamlessly to interpersonal intelligence when comforting a struggling student. AI, by contrast, is often trapped within a single mode of processing—logical and linear, efficient but limited. Gardner's theory compels us to question whether systems like Tutor CoPilot, which excel at pattern recognition and language processing, can ever truly support the full spectrum of human cognitive diversity.

Herbert A. Simon, one of the pioneers of cognitive psychology, argued that expertise is defined not by the possession of knowledge but by the way it is organized and applied in context (Simon, 1973). Expert tutors don't just know more—they see differently. Where a novice sees a jumble of disconnected facts, an expert sees a coherent structure, effortlessly identifying the best path forward. AI systems, despite their formidable computational power, struggle to emulate this kind of perceptual organization. Tutor CoPilot might be able to identify common misconceptions based on data, but it lacks the "feel" for when to press a point or when to let a student grapple with the problem a bit longer.

This gap is what Ben Shneiderman has termed the "human-centered AI dilemma" (Shneiderman, 2020). As AI becomes more sophisticated, it risks overshadowing the very human skills it was designed to support. Shneiderman's research into human-computer interaction suggests that the true potential of AI lies not in replacing human intuition but in augmenting it—offering subtle nudges and cues that enhance human judgment without drowning it out. For Tutor CoPilot, the challenge is finding that balance: becoming a guide that bolsters a tutor's confidence without becoming a crutch.

This delicate interplay between support and overreach is reminiscent of Daniel Susskind's critique of AI's role in the future of professional expertise (Susskind and Susskind, 2015). As AI systems become more capable, they risk simplifying complex professional practices into a series of if-then statements, stripping out the nuances that define true expertise. For a tutor like Emily, this could mean relying too heavily on Tutor CoPilot's suggestions, losing the opportunity to develop her own instructional instincts. Susskind's point is not to deny the value of AI, but to caution against the unintended consequences of embedding it too deeply into processes that rely on human judgment.

Yet, the allure of AI remains strong, precisely because of the unique challenges of scaling expertise in human-centered fields. Erik Brynjolfsson and Andrew McAfee have argued that AI's greatest impact will not be in automating tasks but in complementing human strengths (Brynjolfsson and McAfee, 2014). In education, this means using AI to amplify the skills of novice tutors, allowing them to operate at a higher level than their experience would normally permit. But this raises a critical question: If AI can make a novice appear as skilled as a seasoned professional, what happens to the craft of teaching? Do we lose something irreplaceable in the process?

Thomas Malone, a leading voice in the study of human-AI collaboration, suggests that the solution lies in redefining the relationship between humans and machines, not as competitors but as partners (Malone, 2018). His research on collective intelligence shows that the best outcomes arise when human intuition and AI precision are brought together in a dynamic interplay. In the case of Tutor CoPilot, this means creating a system that doesn't just provide answers but fosters deeper reflection, prompting tutors to think critically about why a particular approach might work better than another.

There is, however, a darker side to this narrative. Shoshana Zuboff's concept of "surveillance capitalism" warns that as AI systems become more ingrained in human activities, they also become powerful tools for control and manipulation (Zuboff, 2019). If Tutor CoPilot can track every decision a tutor makes, every hesitation or mistake, it risks turning a supportive tool into a mechanism of surveillance, subtly shaping behavior to fit predefined models of "good" teaching. This, as Evgeny Morozov argues, is the paradox of digital utopianism: the tools designed to liberate can just as easily constrain (Morozov, 2013).

This tension between potential and peril underscores the need for a nuanced approach to AI integration in education—one that acknowledges both its strengths and its profound limitations. As we move forward, the task is not just to build better AI but to ensure that it serves the complex, deeply human goals that define education. Because, in the end, the true measure of Tutor CoPilot's success won't be how well it mimics human expertise, but how well it enhances it without eroding the unique, irreplaceable qualities that make teaching a profoundly human endeavor.

The Promise and Pitfalls of AI in Educational Settings

In 1956, a group of researchers gathered at Dartmouth College to answer a question: Could machines learn to think like humans? It was a bold question, one that would launch the field of artificial intelligence and reshape the future of human endeavor. But over sixty years later, as we see AI like Stanford's Tutor CoPilot entering the classroom, the question has evolved. Now, it's not just about whether machines can think, but whether they can teach. Can AI offer the same level of nuance, empathy, and understanding that a skilled educator can? The answer, like most things involving human nature, is complicated.

Theoretical Foundations

Theoretical frameworks around teaching and learning have long centered on a basic truth: expertise in education is more than the sum of its parts. Lee Shulman, a foundational figure in educational psychology, coined the term "pedagogical content knowledge" (PCK) to capture what sets expert teachers apart from novices (Shulman, 1987). A teacher isn't just a subject-matter expert; they possess a special kind of knowledge that enables them to present complex concepts in ways that students can grasp. It's this PCK—this intuitive ability to adjust a lesson on the fly, to shift metaphors when one explanation doesn't land—that AI systems like Tutor CoPilot struggle to emulate.

To grasp why AI struggles with something so seemingly simple, we must consider the work of Herbert Simon, one of the earliest cognitive psychologists to study human expertise. Simon argued that expert knowledge is not a static repository of facts; it's a rich, interwoven tapestry of experiences and understandings, built up over years of practice (Simon, 1973). For a seasoned teacher, every lesson is informed by countless encounters with confused faces and "aha" moments, each one adding a new thread to the tapestry. An AI, no matter how advanced, operates outside this tapestry. It can process data and recognize patterns, but it lacks the lived experience that turns information into wisdom.

This discrepancy is why Linda Darling-Hammond has long emphasized the importance of practice in teacher development. Her research reveals that true pedagogical expertise emerges only when theory meets practice—when aspiring teachers confront the unpredictability of real classrooms (Darling-Hammond, 2006). The unpredictability is precisely what AI like Tutor CoPilot cannot fully grasp. When a student's eyes glaze over or their voice wavers with uncertainty, a human teacher instinctively knows to pause, to switch tactics, or to offer a word of encouragement. An AI can flag that a student is struggling based on their quiz scores, but it cannot perceive that struggle in real-time, nor can it discern the emotional undertones that dictate the best response.

Consider Dylan Wiliam's work on formative assessment, which advocates for a responsive, student-centered approach to teaching. Wiliam's theories suggest that the most effective teaching happens when educators constantly adapt their strategies based on nuanced student feedback (Wiliam, 2011). Here lies the rub: Tutor CoPilot, for all its promise, functions best within a narrow band of well-defined parameters. It can suggest alternative explanations for a math problem, but it cannot yet gauge when a student's frustration stems from the problem itself or from the anxiety of feeling "stupid." The art of teaching, as Wiliam reminds us, is an exercise in reading between the lines—something that AI, even the best of it, has yet to master.

But what about the counterargument? Can AI, in its own way, enhance aspects of teaching that even the best human educators struggle with? That's where proponents like Robert J. Marzano enter the picture. Marzano's research on effective classroom strategies shows that certain instructional methods—like scaffolding, spaced repetition, and immediate feedback—consistently produce better learning outcomes (Marzano, 2003). AI, by design, excels at delivering these strategies with precision and consistency. While a human teacher might struggle to track every student's progress in a large classroom, Tutor CoPilot can instantly identify which students need more practice and adjust its recommendations accordingly. It's a tantalizing vision: a teaching assistant that doesn't forget, doesn't get tired, and is always on hand to help.

Yet, for all its strengths, AI's rigid adherence to pre-set strategies highlights a key weakness. As Etienne Wenger would argue, learning is not just about the transfer of knowledge; it's about the formation of *communities of practice* (Wenger, 1998). In classrooms, teachers don't just teach content—they model the social norms and behaviors that shape a learning community. They foster collaboration, nurture curiosity, and help students navigate the complex web of social interactions that define school life. AI, lacking the ability to engage in social contexts, remains a solitary actor in a collective space. It can support individual learning paths, but it cannot yet become a part of the learning community.

AI in Action—Human-AI Collaboration Models

The question, then, is how to integrate AI in ways that complement rather than replace human strengths. One promising approach comes from frameworks developed by researchers like Carolyn Penstein Rosé and Pierre Dillenbourg, who have explored human-AI partnerships in educational contexts (Dillenbourg, 1999; Rosé et al., 2015). Their studies show that AI can play a valuable role as a "cognitive partner," providing suggestions and insights that expand a teacher's toolkit without overshadowing their judgment. In this model, the AI is not the lead instructor but a backstage assistant, enhancing the teacher's performance without stepping into the spotlight.

Stanford's Tutor CoPilot embodies this vision, functioning less as a "super-tutor" and more as a knowledgeable peer. It suggests strategies, flags potential issues, and offers supplemental resources, but always leaves the final decision to the teacher. This model aligns with what Rose Luckin calls the "human plus" approach, where AI augments human capabilities by filling in gaps in knowledge or tracking patterns that might otherwise go unnoticed (Luckin, 2017). In Luckin's view, the value of AI is not in its ability to automate teaching but in its capacity to support and enhance the distinctly human elements of education.

Yet, as we explore these models, we must heed the warnings of critics like Neil Selwyn, who cautions against the uncritical adoption of technology in educational settings (Selwyn, 2011). For Selwyn, the danger lies not in the AI itself but in the assumptions we make about its role. If we view AI as a panacea, we risk designing systems that prioritize efficiency over empathy, data over dialogue. AI can undoubtedly support novice educators like Emily, but it must do so in a way that preserves the art of teaching—the delicate balance of knowledge, intuition, and human connection that no machine, no matter how advanced, can ever fully replicate.

So, as Tutor CoPilot and its counterparts continue to evolve, we are left with a familiar but still pressing question: How do we harness technology's power without losing the humanity at the heart of education? As with all great revolutions, the answer may not lie in the technology itself, but in how we choose to use it.

Analyzing Tutor CoPilot's Effectiveness: Evidence from the Study

It's tempting to think that technology can fix what's broken in education. The story of every new digital tool—from the chalkboard to the smartboard—has been marked by similar promises: better learning, improved outcomes, and, of course, data to back it all up. So when Stanford's Tutor CoPilot reported a 4% improvement in learning outcomes overall and a striking 9% boost for novice tutors, it seemed like one more chapter in the familiar tale of technological salvation. But those numbers are more than just data points. They carry within them the ambitions, limitations, and unexpected twists that define the complex dance between human teaching and machine learning.

Quantitative Results and Learning Outcomes

Stanford's randomized controlled trial (RCT) design was rigorous by any standard. The researchers set out to answer a straightforward question: Could Tutor CoPilot, an AI-powered support system, improve the effectiveness of novice tutors? The study involved a diverse group of educators—some seasoned, some fresh out of their training programs. The results, at first glance, seemed promising. A 4% average increase in student performance may not sound dramatic, but as John Hattie's work on *Visible Learning* reminds us, even modest effect sizes can translate into significant educational gains when scaled across entire systems (Hattie, 2009).

Hattie's extensive meta-analyses show that many common educational interventions—think smaller class sizes or increased funding—often produce disappointing results when evaluated in isolation. In this context, Tutor CoPilot's 4% improvement is comparable to interventions like one-on-one tutoring or formative feedback, both of which are considered among the most effective strategies in Hattie's hierarchy. But there's a catch: the effect size wasn't uniform. The 9% improvement observed for novice tutors was where the real story lay.

This discrepancy brings us to a critical nuance that Eric Hanushek, a scholar renowned for his research on teacher quality, has highlighted: the impact of an educational intervention often depends not just on *what* is being implemented, but *who* it's being implemented with (Hanushek, 2011). Tutor CoPilot didn't just improve learning outcomes overall—it closed the gap between novices and experts. It acted as a stabilizer, narrowing the disparity in instructional quality that so often hampers new educators. In doing so, it offered a glimpse of AI's potential to democratize teaching effectiveness. But does a 9% improvement mean that a novice tutor with Tutor CoPilot is now on par with a seasoned educator? Hardly.

Henry Levin's frameworks on educational cost-effectiveness caution against taking such results at face value (Levin et al., 1981). While a 9% boost might justify the adoption of AI for novice tutors, it raises uncomfortable questions about long-term sustainability. If Tutor CoPilot becomes a crutch, will these same novice educators continue to improve once the AI is removed? In Levin's terms, the challenge is not just about proving impact but about ensuring that the intervention has a durable, compounding effect on teacher expertise. In this light, the trial's findings are more of a starting point than a conclusion.

Qualitative Insights from Tutor Feedback

The numbers tell only part of the story. For a deeper understanding, we have to turn to the people who experienced Tutor CoPilot firsthand—the tutors themselves. And here, the narrative takes on a more complex hue. While many participants reported feeling more confident and prepared, their feedback revealed a subtle tension: some found themselves deferring too much to the AI's suggestions, while others resisted it altogether, feeling that the machine's advice lacked the nuance and situational awareness that only human judgment can provide.

This divide is reminiscent of the dynamic that David Berliner has long described in his studies on teacher expertise (Berliner, 1994). According to Berliner, novice teachers often operate in a state of "instructional triage," struggling to balance content delivery, classroom management, and student engagement simultaneously. AI, by providing real-time support, offers a lifeline. But it's a double-edged sword. If the AI's recommendations are too prescriptive, it can trap novices in a cycle of dependency. If too generic, it becomes background noise—another layer of distraction in an already chaotic environment.

Madeline Hunter's instructional theory sheds light on why this is such a delicate balance (Hunter, 1982). The "Hunter Lesson Plan Model" emphasizes the need for structured support that gradually fades as the teacher gains mastery. Tutor CoPilot, however, operates on a binary scale—it's either on or off. This lack of graduated support might explain why some tutors felt overwhelmed by the AI's constant presence. It's like having a helpful colleague who never stops offering unsolicited advice, even when you just need a moment to think.

For others, the AI's limitations became painfully obvious. Barbara Rogoff's theories on guided participation suggest that effective teaching involves more than just knowledge transmission; it requires the tutor to engage in a shared cognitive journey with their students (Rogoff, 1990). But Tutor CoPilot, being an outsider to the cognitive and emotional rhythms of the classroom, often missed these subtleties. One tutor recounted an experience where the AI suggested a procedural approach to solving a math problem that the student had already grasped conceptually. "It was like getting advice from someone who'd read all the books but never actually taught a class," the tutor remarked.

These narratives complicate the quantitative findings, raising questions about whether Tutor CoPilot is genuinely enhancing tutor expertise or merely filling in gaps in knowledge. Paul Black and Dylan Wiliam's work on formative assessment offers a potential answer (Black & Wiliam, 1998). They argue that the most effective feedback is not just corrective but transformative—it prompts a shift in the learner's understanding. For tutors, this means developing the metacognitive skills to diagnose student needs in real-time and adjust accordingly. But Tutor CoPilot, despite its impressive analytics, often struggled to support this kind of adaptive learning. It could highlight a problem but couldn't explain *why* a particular strategy might work better, leaving tutors in a state of perpetual uncertainty.

Alternate Perspectives: A Critical Lens

The debate over Tutor CoPilot's effectiveness becomes even more charged when we consider alternate voices. Bastani et al. (2024) argue that AI's role in teacher training should be limited to well-defined domains where its guidance is unambiguous and actionable. Their research shows that when AI strays into areas requiring interpretative judgment, such as assessing student engagement or emotional states, it risks undermining the tutor's confidence. This finding resonates with Noam Chomsky's critique of AI's semantic limitations (Chomsky, 2011). Chomsky has long contended that while AI can simulate certain aspects of human language and behavior, it lacks the deep, contextual understanding that comes from lived experience.

In a similar vein, Dietrichson et al. (2017) caution against over-relying on technology in teacher training. Their meta-analysis found that while digital tools can accelerate the acquisition of procedural skills, they often fall short in fostering the reflective practices that define expert teaching. This concern is echoed by Meredith Broussard, whose book *Artificial Unintelligence* critiques the over-optimistic adoption of AI in contexts that require human empathy and nuance (Broussard, 2018). Broussard warns that AI's allure often blinds us to its fundamental limitations, leading to what she calls "techno-solutionism"—the misguided belief that every human problem has a technological fix.

The implications of these critiques are profound. They suggest that Tutor CoPilot's greatest strength—its ability to provide real-time feedback—may also be its greatest weakness. By inserting itself too deeply into the instructional process, it risks disempowering the very people it was designed to support. As Gary Marcus has pointed out, the danger of AI in education is not that it will replace teachers but that it will create a generation of educators who rely on machines to think for them (Marcus, 2020).

In the end, Tutor CoPilot's 4% improvement and 9% boost for novice tutors are more than just data points. They are a testament to both the promise and pitfalls of AI in education. The numbers tell a story of potential—of an AI system that can elevate the skills of novice educators and help level the playing field. But they also hint at a deeper, more troubling narrative: one where the technology designed to enhance human expertise ends up constraining it instead. As we look to the future, the challenge will be finding ways to harness AI's strengths while preserving the uniquely human elements of teaching that no algorithm, no matter how sophisticated, can ever replicate.

Unpacking the Challenges: Where Does Tutor CoPilot Fall Short?

In the fall of 2017, a group of teachers from a diverse range of schools gathered at a conference to witness the latest in educational technology: Stanford's Tutor CoPilot. The excitement was palpable. Here was a tool that promised to revolutionize the classroom by guiding novice tutors through real-time instructional support, offering suggestions, feedback, and insights to help them navigate the complex terrain of teaching. But as the initial demonstrations concluded and the audience began asking questions, the mood shifted. "What happens if the AI misjudges a student's readiness?" asked one experienced teacher. Another wondered aloud: "How does the system handle different grade levels and learning contexts?" The answers, as it turned out, weren't quite as compelling as the initial pitch.

The problem isn't that Tutor CoPilot doesn't work—it's that it doesn't always work where it's needed most. This discrepancy between theoretical promise and real-world application is a common pitfall in the integration of AI across various fields, not just education. And it's a gap that has deep theoretical roots. Take Donald Schön's concept of *reflective practice*, which argues that the value of an expert lies not in their ability to follow predefined rules, but in their capacity to navigate complex, ever-shifting landscapes (Schön, 1983). The real-world classroom, with its unpredictable dynamics and nuanced interactions, is precisely such a landscape. AI, for all its computational prowess, often finds itself out of its depth.

Real-World Applicability vs. Theoretical Promise

One of the key limitations highlighted in the Stanford study is Tutor CoPilot's difficulty in adapting to different grade levels and subjects. This may seem like a minor issue, but it speaks to a larger problem: the lack of contextual understanding. AI, by its very nature, processes information through predefined algorithms and models. While this works well in structured environments, it falters in situations that demand flexibility and nuance. To borrow from Michael Polanyi's famous concept of *tacit knowledge*, there are certain things we know that we cannot easily articulate—things like the tone of a student's voice that signals disengagement or the subtle signs of frustration that suggest a lesson needs to be paused (Polanyi, 1966). Tutor CoPilot, bound by its algorithms, lacks this "feel" for the classroom.

Jean Lave's work on *Situated Learning* further illuminates this challenge (Lave, 1988). Lave argues that learning is inherently tied to context; it cannot be separated from the environment in which it occurs. In traditional classrooms, tutors rely on cues from their surroundings—student reactions, classroom layout, even the time of day—to adjust their teaching strategies. Tutor CoPilot, in contrast, operates in a context-agnostic manner. It delivers feedback and suggestions as if teaching were a static process, devoid of the situational factors that shape every learning moment. It's like trying to dance to music you can't quite hear—the steps might be correct, but the rhythm is off.

This misalignment is reminiscent of what Karl Weick describes as a breakdown in *sensemaking* (Weick, 1995). When human teachers encounter a situation that doesn't fit neatly into their mental model, they adjust, reframe, and, ultimately, find a new way to interpret the situation. AI, lacking this ability, tends to either overcorrect or fail entirely. In the Stanford trial, this played out in scenarios where Tutor CoPilot misinterpreted student responses, suggesting remedial activities that were either too advanced or too simplistic. The result? Confused students and frustrated tutors.

The problem extends beyond just instructional mismatches. Chris Argyris's theories on *organizational learning* highlight how systems that lack the capacity for double-loop learning—where feedback leads to changes in the system's underlying assumptions—are doomed to stagnate (Argyris, 1999). Tutor CoPilot, for all its real-time feedback, operates within a single-loop learning model. It can adapt within a given framework but struggles to recognize when the framework itself needs to change. This rigidity limits its ability to support continuous professional development, one of the key promises of AI in education.

Privacy, Safety, and Ethical Considerations

But perhaps the most pressing concerns aren't about Tutor CoPilot's instructional efficacy at all. They're about privacy and control. In an era when data is often considered the new oil, the implications of allowing an AI to monitor and record every aspect of the tutoring process are profound. Helen Nissenbaum's work on *privacy in context* provides a framework for understanding why (Nissenbaum, 2010). Nissenbaum argues that privacy is not about secrecy but about appropriate flows of information. What happens when an AI system like Tutor CoPilot is privy to every hesitation, every incorrect answer, every pause in a student's response? Does it make students more self-conscious? More guarded? The Stanford study didn't fully address these questions, but the potential for surveillance is clear.

Virginia Eubanks's book *Automating Inequality* sheds light on a darker side of AI's presence in education (Eubanks, 2017). She shows how well-intentioned systems can perpetuate existing disparities by embedding biases into their decision-making processes. Tutor CoPilot, designed to provide equitable support, could just as easily become a tool for reinforcing stereotypes if its algorithms aren't carefully scrutinized. For example, if the system consistently flags students from certain backgrounds as "at risk," it may lead to lower expectations and, ultimately, lower performance—a self-fulfilling prophecy of failure.

This concern is echoed by Cathy O'Neil's critique of algorithmic bias in *Weapons of Math Destruction* (O'Neil, 2016). O'Neil argues that AI systems, because of their opacity, are often immune to accountability. If a tutor follows Tutor CoPilot's advice and the student still struggles, who is to blame—the AI or the teacher? This lack of clarity not only undermines trust but also raises serious ethical questions about responsibility and agency. Brian Christian's recent work, *The Alignment Problem*, delves deeper into these issues, highlighting the risks of AI systems that are misaligned with human values (Christian, 2020). In the case of Tutor CoPilot, the alignment problem manifests in subtle but significant ways—suggesting rote exercises when creative problem-solving is needed or prioritizing test scores over deeper learning.

Contrarian perspectives, such as those from Jaron Lanier and Evgeny Morozov, take these concerns even further. Lanier, a pioneer in virtual reality, has long warned that AI systems risk reducing human creativity and autonomy (Lanier, 2010). In his view, tools like Tutor CoPilot are not just educational aids but potential gatekeepers, subtly shaping how we think and learn. Morozov, in his critique of digital utopianism, argues that the very premise of using AI to solve complex social problems is flawed (Morozov, 2013). Education, he suggests, is not a series of technical challenges to be optimized but a deeply human endeavor that requires empathy, trust, and moral judgment—qualities that no algorithm, no matter how advanced, can replicate.

This sentiment is shared by Sherry Turkle, who has spent decades studying the impact of technology on human relationships. Turkle's research shows that as we outsource more of our thinking to machines, we risk losing the very skills that make us uniquely human—patience, empathy, and the capacity for deep, reflective thought (Turkle, 2015). Nicholas Carr's book, *The Shallows*, expands on this theme, suggesting that our increasing reliance on AI-driven systems may be reshaping our brains, making us less able to engage in the kind of sustained focus and critical thinking that true learning requires (Carr, 2010).

So, where does this leave Tutor CoPilot? It's clear that the system has potential. It can provide real-time support, enhance novice tutors' confidence, and offer data-driven insights that would be impossible for a human to generate on the fly. But it's equally clear that there are limits—both practical and ethical. The real challenge for Tutor CoPilot, and for AI in education more broadly, is to navigate these boundaries without overstepping. AI can augment, but it cannot replace. It can guide, but it cannot lead. And until we fully grasp the implications of what it means to bring machines into the most human of professions, we risk creating tools that do more harm than good. In the end, the success of Tutor CoPilot won't be measured by its effect sizes or adoption rates, but by how well it manages to stay in its place: as a partner, not a master, in the art of teaching.

Comparative Analysis: Tutor CoPilot vs. Other Human-AI Approaches

Let's recall that fresh, new teacher that we first imagined, nervously flipping through her notes, unsure of how to explain a maths problem that half her students grasp and the other half look utterly perplexed by… until Tutor CoPilot prompts her and the fog of confusion lifts from her students' faces. Now imagine a radiologist sits in a dark room, examining a grainy image of a lung scan. AI-powered software, trained on millions of medical images, flags a tiny, almost imperceptible nodule that the radiologist might have missed. "Possible early-stage cancer," the prompt reads. The radiologist leans forward, scrutinizing the highlighted area. It's a life-altering moment, one in which human expertise and machine intelligence converge.

These two scenarios may seem worlds apart—one in the classroom, the other in a hospital—but they share a common thread: AI stepping in to assist human experts at critical decision points. In education, as in medicine, law, and other complex fields, the promise of AI lies not in replacing human judgment but in enhancing it, amplifying our strengths and compensating for our weaknesses. But while AI's potential seems boundless, its pitfalls are equally profound. And nowhere is this tension more visible than in a direct comparison between systems like Tutor CoPilot and AI approaches in other high-stakes domains.

AI in Tutoring vs. AI in Other Domains

Let's start by looking at how AI has made its way into medicine. Eric Topol, a leading expert on AI in healthcare, has long advocated for the use of AI as a diagnostic tool that augments—rather than substitutes—human expertise (Topol, 2019). Take AI-assisted diagnosis: systems like IBM's Watson can scan thousands of medical records, identify patterns invisible to the human eye, and suggest potential diagnoses in seconds. Yet, Topol is careful to remind us that while AI can point out abnormalities, it's still up to the doctor to interpret these results within the broader context of the patient's health. This nuance mirrors the limitations seen in Tutor CoPilot: just as Watson cannot consider a patient's unique emotional state or personal history, Tutor CoPilot cannot fully grasp the complex interpersonal dynamics that play out in a classroom.

Siddhartha Mukherjee, a physician and Pulitzer Prize-winning author, offers another perspective on AI's role in medicine. He describes the interaction between doctor and patient as an art form—a subtle, evolving dance that defies reduction to data points and probabilities (Mukherjee, 2016). When AI systems like Tutor CoPilot step into the educational arena, they confront a similar challenge. Education, like medicine, is as much about human connection as it is about transferring knowledge. An AI tutor may be able to provide step-by-step problem-solving guidance, but it cannot read the worried frown of a student who is struggling to keep up or sense the triumph of a student who has finally grasped a difficult concept.

In the legal world, Richard Susskind has written extensively about how AI is transforming the practice of law through tools like AI-guided legal research systems (Susskind, 2017). These systems can sift through thousands of case files in moments, pinpointing relevant precedents and highlighting potential weaknesses in an argument. The result is a more efficient legal process, but one that still relies on the lawyer's skill to craft a compelling narrative. Tutor CoPilot operates similarly: it offers suggestions and resources, but it's up to the human tutor to weave these into a cohesive lesson that resonates with students. Cass Sunstein's research on decision-making in regulatory contexts further underscores this point. AI can guide and inform, but ultimately, it is the human's role to interpret, contextualize, and decide (Sunstein, 2019).

Yet, there are crucial differences between education and these other fields. In healthcare and law, AI operates within well-defined domains with clear rules and outcomes. In contrast, the classroom is a fluid, ever-changing environment. This is where Luciano Floridi's concept of *AI as a moral agent* comes into play (Floridi, 2013). Floridi argues that when AI systems begin to influence human behavior—whether in legal advice or educational guidance—they assume a form of agency. In the classroom, this agency is fraught with ethical complexities. What happens when Tutor CoPilot's guidance contradicts a tutor's intuition? Whose judgment should prevail?

Innovations in Adaptive Learning—Where Does Tutor CoPilot Stand?

Tutor CoPilot's promise lies in its ability to adapt to the needs of individual students in real-time, placing it within a broader landscape of AI-driven educational platforms like Khan Academy and OpenAI's Codex. But does it stand out? Arthur Graesser, a pioneer in intelligent tutoring systems, might say it depends on how you define "adaptive" (Graesser et al., 2001). His work on AutoTutor, an AI-based system that engages students in dialogue, emphasizes that true adaptivity involves more than just responding to a student's errors. It requires the system to understand *why* the student is struggling and adjust its strategies accordingly.

Neil Heffernan's ASSISTments platform offers another point of comparison. Like Tutor CoPilot, ASSISTments provides real-time feedback to both students and teachers. But Heffernan's approach prioritizes transparency and teacher control. Rather than offering prescriptive solutions, ASSISTments presents data in a way that empowers educators to make informed decisions (Heffernan et al., 2014). Tutor CoPilot, by contrast, risks overstepping its bounds by offering direct guidance, potentially undermining the tutor's autonomy. As Peter S. Rosenbloom, developer of Andes Physics Tutor, has pointed out, the key challenge for any AI tutor is not to become the "teacher," but to remain a supportive partner in the learning process (Rosenbloom, 2006).

The issue becomes even more pronounced when we consider Carol Dweck's research on *Growth Mindset* (Dweck, 2006). Dweck's theories suggest that the language we use to give feedback shapes how students view their own abilities. If Tutor CoPilot's suggestions are framed too rigidly—focusing on right or wrong answers rather than effort and strategy—it could inadvertently reinforce a fixed mindset. The very tool designed to foster learning might, if poorly implemented, end up stifling it.

But for all these concerns, Tutor CoPilot has one distinct advantage over its competitors: its integration of real-time instructional support. While most AI platforms focus on student learning, Tutor CoPilot shifts the focus to the educator, offering guidance and resources in the moment they're needed most. Tom Davenport and H. James Wilson, authors of *Human + Machine*, argue that the most successful AI systems are those that amplify human strengths rather than automate them (Davenport & Wilson, 2018). By supporting tutors rather than supplanting them, Tutor CoPilot aligns with this vision.

The debate over AI in education, as Diane Ravitch warns, is often clouded by the allure of efficiency (Ravitch, 2010). Systems like Tutor CoPilot are celebrated for their potential to streamline instruction and boost outcomes. But as Andrew Ng has noted, the true potential of AI lies not in making humans more efficient but in making us more effective (Ng, 2016). For Tutor CoPilot, this means moving beyond just optimizing instructional time and focusing on empowering tutors to become better, more reflective educators.

In the end, Tutor CoPilot stands at a crossroads. It is a unique experiment in leveraging AI to support—not replace—human expertise. Compared to its counterparts in medicine and law, it faces a more complex, less predictable environment. The classroom is not a courtroom or an operating theater; it is a living ecosystem where every interaction shapes the next. To thrive in this space, Tutor CoPilot must navigate a delicate balance: augmenting human intuition without overshadowing it, offering guidance without taking the reins. Whether it succeeds will depend not on the brilliance of its algorithms, but on the wisdom of its designers and the educators who choose to embrace it. Because in the end, the real promise of AI in education is not about creating better machines; it's about creating better humans.

The Future of Expertise in the Age of AI

It's hard to imagine now, but there was a time when the introduction of calculators into classrooms sparked outrage. Teachers worried that students would stop learning basic arithmetic, parents feared the devices would undermine education, and policymakers debated whether technology was a crutch or a bridge to new ways of thinking. Fast forward a few decades, and calculators are just another tool in the educational arsenal, a mundane fixture on every student's desk. The lesson is clear: disruptive technologies often provoke anxiety not because they change how we teach, but because they challenge us to rethink what it means to be educated.

Tutor CoPilot is our modern-day calculator, but with far greater stakes. Where the calculator enhanced our ability to perform computations, Tutor CoPilot promises to transform how we acquire, scale, and distribute human expertise. But as we've seen, this promise is fraught with complications. AI, in education and beyond, often evokes a tug-of-war between utopian aspirations and very real limitations. What if, instead of choosing sides, we asked a different question: What should we *want* AI to do for us? Because ultimately, Tutor CoPilot isn't just a tool—it's a prototype for a new kind of partnership between humans and machines.

In this article, we've explored Tutor CoPilot's strengths and shortcomings through many lenses: pedagogy, cognitive science, ethics, and policy. The broader narrative that emerges is not one of success or failure, but of potential and purpose. When Tutor CoPilot helps a novice tutor navigate a tricky lesson, it's not replacing the teacher—it's amplifying their potential. And when it misses the mark, suggesting rote solutions to complex problems, it reveals not just the limitations of technology, but also the vast, uncharted territory of human judgment that machines struggle to map.

This brings us to the deeper question that thinkers like Ray Kurzweil have been grappling with for decades: How do we design technologies that don't just mimic human expertise, but *enhance* what makes us uniquely human? Kurzweil envisions a future where humans and machines are collaborators, each complementing the other's strengths to achieve things neither could alone (Kurzweil, 2005). Tutor CoPilot, in its current form, is a tentative step toward this vision—a system that hints at what's possible but also lays bare the gaps that still need to be filled.

Max Tegmark, in his work on the future of artificial intelligence, pushes this idea further. He argues that as machines become more capable, the challenge is not just to make them more like us, but to make sure they are aligned with our deepest values and aspirations (Tegmark, 2017). In education, this means designing AI systems that support curiosity, empathy, and resilience—traits that define good teachers and lifelong learners. Tutor CoPilot, with its focus on real-time support and formative feedback, edges toward this ideal. But it must evolve beyond being a suggestion engine; it must become a catalyst for deeper engagement and reflective practice.

Yet, as Yuval Noah Harari reminds us, the integration of AI into education isn't just about technology—it's about power (Harari, 2018). AI systems, whether in classrooms or courtrooms, shape not only how decisions are made but *who* makes them. This is why it's crucial to ensure that Tutor CoPilot remains a tool that empowers teachers rather than disempowers them, that it augments human insight rather than constrains it. If we let the convenience of AI obscure the need for human judgment, we risk creating a generation of educators who defer to machines rather than honing their own instincts.

Ken Robinson's theories on creativity and divergent thinking add another layer to this dilemma (Robinson, 2006). Education is not just about transmitting knowledge; it's about cultivating the capacity for innovation and creative problem-solving. The danger with AI systems like Tutor CoPilot is that they might inadvertently prioritize efficiency over creativity, standardizing instructional methods in ways that stifle the very diversity of thought they are meant to nurture. The real challenge, then, is to design AI that celebrates the unpredictable nature of human learning rather than trying to eliminate it.

As we look to the future, perhaps the most provocative insight comes from Hubert Dreyfus, a philosopher who has spent his career critiquing the overreach of AI (Dreyfus, 1972). Dreyfus argued that human expertise is deeply embodied, shaped by experience, intuition, and a sense of timing that machines cannot replicate. Tutor CoPilot, no matter how advanced its algorithms, cannot feel the pulse of a classroom, cannot sense the subtle shift in a student's attention, cannot grasp the silence that signals a breakthrough moment. It can prompt, suggest, and guide—but it cannot *teach* in the way that great educators do.

But does this mean we should dismiss Tutor CoPilot as a failed experiment? Not at all. It means that we should view it for what it is: a prototype, a work in progress, a stepping stone toward a new kind of educational technology that doesn't just replicate human expertise but transforms it. And to do that, we must reimagine the role of AI in education from the ground up. Esther Wojcicki's *Moonshots in Education* approach offers a compelling framework here (Wojcicki, 2015). Wojcicki believes that the goal of educational technology should be to inspire students and teachers to dream bigger, to reach for goals that would be impossible without technology. Tutor CoPilot, in its best moments, has this potential—it can open new pathways for learning, create spaces for reflection, and provide support where it's needed most.

So where do we go from here? The answer lies in embracing the paradoxes that AI forces us to confront. We must design systems that are powerful yet humble, systems that push the boundaries of what's possible without overstepping their place in the human learning journey. This is where Donna Haraway's *Cyborg Theory* becomes particularly resonant (Haraway, 1985). Haraway envisions a future where humans and machines are not adversaries but co-evolutionary partners, each shaping the other in profound ways. If Tutor CoPilot can become this kind of partner—if it can support not just instruction but the entire ecosystem of learning—then it will have truly succeeded.

As we close, let's return to our novice tutor, standing nervously at the front of her classroom. She glances at Tutor CoPilot's screen one last time before stepping forward, deciding to trust her own instincts over the AI's advice. The students are watching, waiting. She takes a deep breath, improvises, and sees the light of understanding flicker in their eyes. That's the moment when technology steps back, and human expertise steps forward. It's not about replacing the teacher. It's about helping her find her own voice. Because in the end, the future of education isn't just about machines that teach. It's about building systems that help teachers—and students—discover the full extent of their humanity.

Bibliography

Argyris, C. (1999). *On Organizational Learning*. 2nd ed. Blackwell Business.

Berliner, D. C. (1994). *Expertise: The wonder of exemplary performances*. In J. N. Mangieri & C. C. Block (Eds.), *Creating powerful thinking in teachers and students* (pp. 161–186). Holt, Rinehart & Winston.

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. Phi Delta Kappan.

Broussard, M. (2018). *Artificial Unintelligence: How Computers Misunderstand the World*. MIT Press.

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company.

Carr, N. (2010). *The Shallows: What the Internet Is Doing to Our Brains*. W. W. Norton & Company.

Chomsky, N. (2011). *The Science of Language: Interviews with James McGilvray*. Cambridge University Press.

Christian, B. (2020). *The Alignment Problem: Machine Learning and Human Values*. W. W. Norton & Company.

Davenport, T. H., & Wilson, H. J. (2018). *Human + Machine: Reimagining Work in the Age of AI*. Harvard Business Review Press.

Dede, C. (2008). *A Seismic Shift in Epistemology*. Educause Review, 43(3), 80-81.

Dietrichson, J., et al. (2017). *The effectiveness of educational interventions targeted at students with low socio-economic status: A systematic review and meta-analysis*. Review of Educational Research, 87(2), 243-282.

Dreyfus, H. L. (1972). *What Computers Still Can't Do: A Critique of Artificial Reason*. Harper & Row.

Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. Random House.

Eubanks, V. (2017). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.

Floridi, L. (2013). *The Ethics of Information*. Oxford University Press.

Fullan, M. (2007). *The New Meaning of Educational Change*. 4th ed. Teachers College Press.

Gardner, H. (1983). *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books.

Giroux, H. A. (2011). *On Critical Pedagogy*. Bloomsbury Academic.

Graesser, A. C., et al. (2001). *AutoTutor: A Simulation of a Human Tutor*. Cognitive Systems Research, 1(1), 35-51.

Hargreaves, A. (2005). *Sustainable Leadership*. Jossey-Bass.

Harari, Y. N. (2018). *21 Lessons for the 21st Century*. Random House.

Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Routledge.

Heffernan, N. (2014). *ASSISTments: Helping Teachers Make Data-Driven Decisions*. Technology, Knowledge and Learning, 19(3), 123-144.

Hunter, M. (1982). *Mastery Teaching*. Corwin.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Viking.

Lanier, J. (2010). *You Are Not a Gadget: A Manifesto*. Alfred A. Knopf.

Lave, J. (1988). *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. Cambridge University Press.

Levin, H. M., et al. (1981). *Cost-Effectiveness Analysis in Education*. Sage Publications.

Luckin, R. (2017). *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*. UCL IOE Press.

Malone, T. W. (2018). *Superminds: The Surprising Power of People and Computers Thinking Together*. Little, Brown and Company.

Marzano, R. J. (2003). *What Works in Schools: Translating Research into Action*. ASCD.

Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. PublicAffairs.
Mukherjee, S. (2016). *The Gene: An Intimate History*. Scribner.

Ng, A. (2016). *The Future of Artificial Intelligence*. Stanford University Press.

Nissenbaum, H. (2010). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing.

Perkins, D. (1992). *Smart Schools: From Training Memories to Educating Minds*. Free Press.

Polanyi, M. (1966). *The Tacit Dimension*. Doubleday & Co.

Robinson, K. (2006). *Out of Our Minds: Learning to Be Creative*. Capstone.

Rogoff, B. (1990). *Apprenticeship in Thinking: Cognitive Development in Social Context*. Oxford University Press.

Rosenbloom, P. S. (2006). *A Cognitive Architecture for Education*. AI Magazine, 27(2), 33-44.

Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.

Selwyn, N. (2011). *Education and Technology: Key Issues and Debates*. Bloomsbury Academic.

Shneiderman, B. (2020). *Human-Centered AI*. Oxford University Press.

Shulman, L. S. (1987). *Knowledge and Teaching: Foundations of the New Reform*. Harvard Educational Review, 57(1), 1-22.

Simon, H. A. (1973). *The Structure of Ill-Structured Problems*. Artificial Intelligence, 4(3-4), 181-201.

Strauss, V. (2022). *Education Reimagined*. Harvard Educational Review, 91(1), 123-142.

Sunstein, C. R. (2019). *On Freedom*. Princeton University Press.

Susskind, R., & Susskind, D. (2015). *The Future of the Professions: How Technology Will Transform the Work of Human Experts*. Oxford University Press.

Susskind, R. (2017). *Tomorrow's Lawyers: An Introduction to Your Future*. 2nd ed. Oxford University Press.

Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence*. Knopf.

Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.

Turkle, S. (2015). *Reclaiming Conversation: The Power of Talk in a Digital Age*. Penguin Press.

Tyack, D., & Cuban, L. (1995). *Tinkering Toward Utopia: A Century of Public School Reform*. Harvard University Press.

Wang, R., Ribeiro, A.T., Robinson, C.D., Loeb, S. and Demszky, D. (2024), Tutor Co-Pilot: A Human-AI Approach for Scaling Real Time Expertise, Stanford University Press.

Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.

Wiliam, D. (2011). *Embedded Formative Assessment*. Solution Tree Press.

Wojcicki, E. (2015). *Moonshots in Education: Launching Blended Learning in the Classroom*. Pacific Research Institute.

Zhao, Y. (2009). *Catching Up or Leading the Way: American Education in the Age of Globalization*. ASCD.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.