AI And Machine Learning

A Sports Analogy for Understanding Different Ways to Use AI

by Jake M. Hofman, Daniel G. Goldstein and David M. Rothschild

December 4, 2023

Holloway/Getty Images

Summary.   The potential impact of generative AI on the economy, society, and work is polarizing, swinging from the positive benefits of a technological revolution to doomsday scenarios. The authors have come to think about this issue as points on a spectrum and have... **more**

As with many new technologies, the potential impact of generative AI on society is polarizing. There is a fear that this technology will replace human workers, leading to mass de-skilling and widespread unemployment. And there is a hope that generative AI will instead supercharge human capabilities, leading to unprecedented boosts in individual and collective innovation and productivity. In our work at Microsoft Research on augmenting human cognition and decision-making with AI, we have come to think about these scenarios as points on a spectrum and have created a sports analogy to help think about it: AI tools can range from steroids, to sneakers, to a coach, each representing a different relationship between human users and the technology.

Let's start with steroids, the least desirable point on this spectrum. Steroids elevate short-term performance but leave you

worse off in the long term. For example, imagine a student who is assigned homework to read and critique a passage of text. The student could use a large language model (LLM) as a steroid by simply dropping the text of the passage into ChatGPT or any similar tool, prompting it to produce such a critique, and passing the output back to the teacher. While this gives the student the ability to complete the assignment in an instant, in the long term it defeats the purpose of the exercise. Not only might the student fail to learn from the assignment, but they might never acquire (or slowly lose) the ability to critique a passage and articulate their thoughts in clear writing. Just as with sports, taking shortcuts to achieve a short-term goal can have negative consequences for other goals in the long term.

In the middle of the spectrum, AI-powered tools can instead be used to augment people's skills and make them more productive. That's what we think of as the analog to a good running sneaker. Like Nike's carbon-soled shoes that make runners on average 4-5% faster, some AI tools can elevate performance in the moment without the long-term negative effects of de-skilling workers. For example, LLMs are particularly useful for translating, reformatting, and annotating unstructured text, which can save knowledge workers considerable time and effort. Imagine an analyst who runs an international survey with free-form text responses in 20 different languages. An LLM could drastically reduce the time it takes to translate these responses to one common language, score them for positive or negative sentiment, and extract key themes. Similar to performance-enhancing sneakers, in this case AI can accelerate an analyst's capabilities in the moment, while still allowing for their human judgment, curiosity, and creativity to be applied.

On the most desirable end of the spectrum, AI-powered tools can be used like a coach that improves people's own capabilities. Consider an IT professional who is studying to pass a certification test for a new, complex system. An LLM that has been trained on the system's documentation could interactively generate practice questions as the professional gains competence and provide personalized feedback as it teaches advanced concepts. Ultimately this would help the professional stretch their understanding of the subject matter and develop new skills that last beyond just the time spent using the AI-powered tutor.

Our own research has shown similar benefits for learning new math concepts. In a randomized experiment we ran online, we found that giving people access to LLM-based explanations while practicing standardized test problems led to sizable improvements in their ability to later solve similar test problems on their own. Much like a sports coach, AI tools that teach could

potentially provide both short and long-term benefits to performance.

While we think these parallels between athletic and cognitive capabilities are apt, there are important differences between the two domains. Whereas in sports the three relationships of steroids, sneakers, and coaches are all quite distinct, they are much less so when it comes to AI-powered tools, many of which are small variations on the same underlying technology. As a result, the choices we make in terms of how we design and use these tools can shift them from one end of the spectrum to the other. Likewise, the norms we develop and how our priorities evolve over time will play a crucial role in determining how, when, and why we use AI-powered tools.

Design Choices Matter

First, from a design perspective, seemingly innocuous choices of how we architect AI-powered tools can have a substantial impact on their effects. Take the somewhat mundane example of spell check. Most spell checkers don't automatically fix words as you type without feedback. Instead, they provide cues that indicate a potential misspelling and offer an opportunity to not only correct what's written, but to visualize the proper spelling. This small choice shifts spell check from being a kind of a steroid to more of a coach. It improves your final product while also helping you learn from your mistake (or typo).

By thinking intentionally about how we design AI tools, we can minimize or avoid harmful long-term effects. For instance, our own recent experiments show that simple confidence-based highlighting, similar to what's done in spelling and grammar check, can help people spot and correct fabrications or "hallucinations" produced by LLM-based search tools. The idea is that if we can show people which bits of information in an LLM-generated response might be less reliable, we can help them spot and fix potential errors. This would still offer the chance for productivity gains while maintaining the necessary cues for cognitive awareness of where LLM-based responses can go wrong. With thoughtful design we can develop helpful co-pilots that augment, rather than replace, people in getting work done.

Norms Will Emerge

Second, we will need to develop rubrics for thinking about when and how we use AI-powered tools. Calculators serve as another historical analog here. While we would be hard pressed to make an argument against a banker using a calculator to compute compound interest, it would be equally difficult to make a case for giving a calculator to a grade school student learning basic addition. Just as we have come to settle on these norms for navigating how and when we use past technological innovations, we imagine the same will emerge for AI-powered tools: the same tool may be considered beneficial in some settings but detrimental in others. We also expect these norms to shift over time.

For instance, it used to be the case that physical libraries were the gold standard for finding information and citations, but as the quality and coverage of search engines improved along with our ability to issue effective queries and find the right results, the web has become an increasingly reliable and accepted information source. We believe the same types of shifts will happen with AI-powered tools as the underlying technology and our aptitude for using it co-evolves. At the same time, businesses need to think about the consequences of having workers become overly reliant on the latest tool: having only one way to solve a problem is a point of vulnerability. In the workplace it will fall on business leaders to set norms and expectations for when AI-powered tools are (and aren't) appropriate, and to ensure that they are used responsibly and ethically.

Priorities Will Evolve

None of these categorizations are straightforward and can carry nuance from person to person and industry to industry. Consider how the value placed on spelling and arithmetic has changed over time. As mentioned above, it used to be the case that being able carry out long division or spell obscure words without aid were prized and respected talents, but now they are now largely seen as archaic and esoteric skills. Presumably this is because most people have access to a reliable calculator or spell-checking tool when they need one, transforming what was at some point feared to be a steroid into less of a concern. As a result, this has freed up time for people to invest in developing other skills. In short, we have decided to let some skills atrophy so that we can focus on developing others.

With the growth of AI-powered tools, we expect to see similar shifts in the skills that are valued and prioritized by society. People will use their time differently and invest their learning in new places. For example, detailed syntactical knowledge of specific programming languages might become less valued as generative AI tools evolve to automatically generate code with increasing accuracy, freeing engineers to focus on higher-level system design and architecture. As certain skills become less prominent and new ones emerge as priorities, companies should proactively offer training to ensure employees remain equipped with the most relevant and valued skills.

We hope the steroids, sneakers, coach framework we have provided here will prove useful for thinking through these issues and asking critical questions about the AI tools being woven into both work and life. As we integrate AI, our challenge is to craft tools that enhance rather than diminish human capabilities. Through mindful design and thoughtful consideration of when and how these tools are applied, and allowing these priorities to shift as we co-evolve with them, we have the opportunity to shape a future in which AI augments our human capabilities.

JH

Jake M. Hofman is a Senior Principal Researcher at Microsoft Research in New York City.

DG

Daniel G. Goldstein is a Senior Principal Research Manager at Microsoft Research in New York City.

DR

David M. Rothschild is a Senior Principal Researcher at Microsoft Research in New York City.

X f in 🔖

Read more on **AI and machine learning** or related topics **Developing employees**, **Innovation**, **Organizational learning**, **Technology and analytics** and **Generative AI**