

Beyond Prompt Engineering: What *Co-Intelligence* Offers Education and Leadership

Kshitij Taluja

Undergraduate Researcher, Bachelor of Computer Science, Golden Gate University

Dr. Smrite Goudhaman

Global Professor of Practice, Golden Gate University

Author: Ethan Mollick

Title: *Co-Intelligence: Living and Working with AI*

Publisher: Portfolio / Penguin

Year: 2024

ISBN: 978-0-593-71671-7

Pages: 256

Abstract

Co-Intelligence: Living and Working With AI by Ethan Mollick discusses how generative artificial intelligence (AI) can serve as a "capable, yet imperfect" co-laborer, in addition to being seen as a substitute for human judgment. While this reviewer agrees with many aspects of Mollick's four guidelines, which provide useful direction for user-level practitioners, the review identifies several institutional barriers to implementing these recommendations on a larger scale. The review considers the book from the vantage points of higher education, leadership, governance, costs associated with implementation, educational equity; and the reality of the Global South as well as the labor required to support the use of these technologies. In general, Co-Intelligence provides an excellent introduction to thinking about the ethical considerations involved in adopting AI responsibly. However, there are other organizational, policy, and equity related issues that readers will also need to consider beyond those addressed within the confines of this book.

Keywords: generative AI; higher education; AI governance; educational equity; Global South

Ethan Mollick's "Co-Intelligence: Living and Working with AI" enters an area of the public debate on Generative AI where two primary, extreme views have been expressed – the Evangelists and the Dreaders. The book's primary goal is to present a practical way that both individuals and organizations can use generative AI while still being aware of its potential weaknesses and dangers.

The primary contribution made by Mollick is to think of Large Language Models (LLMs) not as software systems but as untrustworthy co-workers. He develops four rules for developing

"co-intelligence": "invite the AI to the table", "keep a person involved", "tell the model what persona to adopt", and "assume today's system will be the weakest you'll ever experience."

Mollick does his best job at illustrating how one individual might work with these tools, and at pushing the conversation past the idea of simply optimizing prompt strategies. He argues that competency lies in other places, including finding ways to coordinate work over the "jagged frontier" of unpredictable strengths and weaknesses; but more so in knowing when to rely upon it.

Treating AI like a person provides a useful heuristic. It can create a high degree of trust, because a system that speaks like our colleague tends to be believed.

The areas where the ambitions of the author outstrip the evidence include the educational chapters. In "AI as a Tutor," Mollick returns to Benjamin Bloom's "two sigma" problem (Bloom, 1984). Mollick suggests that using AI systems such as Khanmigo, Bloom's results could be achieved at a large scale. The studies done by Bloom were limited to small samples and have not yet been replicated at that magnitude. The ability of teachers to use AI effectively in classrooms will depend upon changing the way student knowledge is assessed, the willingness of faculty members to adjust their teaching methods, and the level of digital literacy among their students. As UNESCO has noted in its guidance on generative AI in education (Miao & Holmes, 2023), these conditions are rarely met. Larger public universities which serve many language populations with limited staffing cannot meet these conditions at the same time; thus Mollick's optimism is localized to specific departments.

The author's conclusions and recommendations are based on the belief that all educational institutions have the capability to independently effect reforms. As such, Mollick's views are a representation of an elite perspective.

The same constraints can exist outside of education. Mollick's model can be applied to other areas of study — including business, government and non-profit organizations. For example, a non-profit entity like Khan Academy could pilot "Khanmigo" due to its ability to secure philanthropic financing; however, a small community organization or a governmental agency could experience many of the same challenges.

Mollick correctly identifies some of the primary issues associated with the adoption of AI, namely the possibility of bias and the limited nature of human oversight. In addition to being transparent about which models are being utilized, there needs to be a clearly defined process

for attributing accountability when systems produce adverse results (OECD, 2019). The reference to Sustainable Development Goal 16 (United Nations, 2015) offers a foundation for accountability but does not provide a method for how collective actions can result in accountable institutions.

Inviting AI into the planning room indicates a willingness to adopt organizational change. Leaders need to do more than simply implement AI in their education practices. They will also need to identify areas where AI will not be used, who will be responsible for the costs associated with AI-related negative outcomes, and how accountability will still exist if both humans and machines work together to measure performance.

A major flaw of this book is that it fails to clearly articulate the true financial implications of using AI responsibly. A critical distinction exists between one individual educator utilizing a chatbot versus an entire institution deploying the same technology. This includes providing ongoing training to staff members, creating policy to define what uses are allowable, allocating funds for licenses, revisiting the organization's data protection protocols, and implementing organizational change management initiatives.

Innovation and building capacity are encouraged through SDG 9, while SDG 10 warns against expanding disparities (United Nations, 2015). Aggressive implementation of educational AI could further widen existing inequalities.

Mollick supports his argument that AI decreases differences in performance levels of employees, based on studies he discusses, including one he co-authored involving consultants (Dell'Acqua et al., 2023) and another examining professional writers (Noy & Zhang, 2023). While this may hold some merit within a single company employing knowledge workers, it falls far short of addressing Mollick's intended goals regarding issues of equity and fairness across all people. Levelling employees' performance levels does not address equitable working conditions and fair compensation for laborers in Kenya, India, and the Philippines, whose labour enables these systems. Questions include: Do the models function in languages other than English? Will the interfaces allow for user accessibility? Are the applications capable of running under low-bandwidth connection and/or minimum hardware requirements? Finally, are the most effective applications locked behind paywalls that restrict access from much of the Global South.

Mollick advises readers to rely on their own judgement, delegate, and have discipline in questioning assumptions. There is considerable distance between what technology can do and

an organization's capacity to take advantage of those technologies; and that it takes time, money, and controversy to close that gap. It should provide useful guidance for higher education administrators, policy makers, and other decision-makers in the private or non-profit sectors who wish to integrate Generative AI into their organizations but maintain human oversight. If read as a single thoughtfully expressed viewpoint (i.e. a western practitioner working in an elite institution) versus a comprehensive guidebook, it would be an excellent place to begin thinking about how we will use and live with AI.

References

- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16.
<https://doi.org/10.3102/0013189X013006004>
- Dell'Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Kraye, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality* (Harvard Business School Working Paper No. 24-013). <https://doi.org/10.2139/ssrn.4573321>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio/Penguin.
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192.
<https://doi.org/10.1126/science.adh2586>
- OECD. (2019). *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449). <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development* (A/RES/70/1). <https://sdgs.un.org/2030agenda>