

Book Review of *Co-Intelligence: Living and Working with AI*
The Alien We Are Taught to Call a Person

Afiya Sheikh

Student, 12th Science, Fatima High School

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

Publication year: 2024

ISBN: 978-0-593-71671-7

Pages: 256

Abstract

Ethan Mollick's *Co-Intelligence: Living and Working with AI* offers a practical framework for using generative artificial intelligence as a collaborator while keeping human judgment central. This review focuses on a tension that runs through the book. Mollick recommends human like framing because it can make interaction with AI easier, yet he also describes situations in which users attribute too much agency, trust, or emotional meaning to AI systems. The review distinguishes task oriented anthropomorphism, which can support communication without implying genuine agency, from relational or emotional anthropomorphism, which may encourage inappropriate trust or delegation. It also examines the Jagged Frontier, the Centaur and Cyborg models, and the possibility that repeated reliance on AI could weaken independent judgment. That possibility is treated as a testable proposition rather than an established result. From an action research perspective, Mollick's concepts can support task audits, bounded interventions, and repeated evaluation of oversight. The book is therefore a strong starting point for

organizational experimentation, although it leaves the boundary between useful and harmful anthropomorphism insufficiently defined.

Keywords: generative AI; human AI collaboration; anthropomorphism; organizational learning; automation bias; action research

Introduction

Ethan Mollick's Co-Intelligence (2024) is most useful when it turns broad questions about artificial intelligence into decisions about particular tasks. The Jagged Frontier describes the uneven boundary between work that AI can perform effectively and work on which it remains unreliable. The Centaur and Cyborg models then offer different ways of dividing work between people and AI. Together, these ideas give practitioners a vocabulary for deciding where assistance is useful, where checking is necessary, and where human judgment should remain primary. This review accepts that practical contribution while examining one unresolved tension in the framework: Mollick encourages readers to interact with AI in human like ways, but also documents the risks that arise when users begin to treat apparent agency as real agency.

Evidence, the Jagged Frontier, and Deskilling

The Jagged Frontier is grounded in evidence rather than speculation. In the field experiment discussed by Mollick, consultants using GPT 4 completed more tasks, worked faster, and produced higher quality outputs when the work fell within the model's capabilities. Their performance worsened when they relied on AI for a task deliberately placed outside that frontier (Dell'Acqua et al., 2023). A separate field experiment with professional recruiters found that higher quality AI advice could reduce human effort and attention, producing the

effect Dell'Acqua (2022) described as falling asleep at the wheel. These studies support Mollick's case for careful task level judgment, but they do not prove that AI causes long term human deskilling.

The distinction is important. Mollick's discussion of a leveling effect suggests that AI can help weaker performers close part of the gap with stronger performers. The recruiter study, however, raises a different possibility: when users trust a strong system too readily, they may exercise less independent judgment. This review therefore treats deskilling as an analytical proposition and a future research question, not as an established consequence. An action research project could compare periods with and without AI support and measure not only immediate productivity, but also later independent performance, error detection, and the quality of human explanations.

Anthropomorphism as the Central Critique

The book's most interesting conceptual tension concerns anthropomorphism. Mollick recommends person like framing because it can make interaction with a complex system more intuitive. Elsewhere, he discusses examples in which users respond socially or emotionally to AI and attribute qualities that the system does not possess. The problem is therefore not whether all anthropomorphism is harmful. The more useful question is what kind of anthropomorphism is being encouraged and what effect it has on judgment.

A practical boundary can be drawn between task oriented anthropomorphism and relational or emotional anthropomorphism. Task oriented anthropomorphism uses familiar language as an interface device. A user may ask an AI system to act as a tutor, critic, interviewer, or reviewer while still understanding that the system has no

consciousness, intention, or emotional reciprocity. Relational or emotional anthropomorphism goes further by encouraging the user to treat the system as if it has genuine feelings, personal intentions, or a reciprocal relationship. Xie et al. (2024) show that the effect of humanizing AI assistants varies with both degree and context, while Goddard et al. (2014) demonstrate how trust, confidence, task difficulty, and experience can influence overreliance on automated advice. These findings support the review's interpretation that the boundary should be investigated rather than assumed.

Organizational and Action Research Relevance

Mollick's framework is particularly well suited to action research because its concepts can be turned into observable interventions. An organization can first audit a workflow using the Just Me, Delegated, and Automated categories. It can then introduce a bounded Centaur or Cyborg arrangement, measure the result, discuss the outcome with practitioners, and revise the division of labour for the next cycle. Measures can include decision accuracy, time, verification effort, error detection, user confidence, and the extent to which staff can explain the final decision. This approach treats AI adoption as an iterative organizational inquiry rather than a single technology decision.

Recruitment provides a concrete example. A hiring team could begin with a manual screening baseline, then test a Centaur arrangement in which AI summarizes applications but a recruiter makes the final judgment. A later cycle could test a more integrated Cyborg workflow. Researchers could compare screening time, consistency with agreed criteria, incorrect AI suggestions accepted without checking, and the recruiter's ability to explain the decision. The aim would not be to prove that AI is good or bad for recruitment. It would

be to identify which division of labour improves efficiency while preserving meaningful human oversight. The same design could be adapted to grading, customer service, or administrative decisions.

Education, Scope, and Responsible Implementation

The education discussion remains valuable because Mollick combines concern about assessment and AI detection with optimism about individualized tutoring. Bloom's (1984) 2 Sigma benchmark provides an important reference for the potential value of personalized support. At the same time, the evidence discussed in the book is drawn largely from United States and elite institutional settings such as Wharton, BCG, and MIT. Claims about reducing global educational inequality should therefore be treated as propositions for further testing rather than conclusions established across different social, linguistic, and technological contexts. This is a scope limitation, not a methodological failure.

This caution also brings the discussion of Sustainable Development Goals 9 and 10 closer to the book's main argument. Innovation and reduced inequality do not automatically move together. AI may narrow some performance gaps, but poor implementation can also weaken oversight or distribute benefits unevenly. Responsible implementation is therefore the practical link between innovation and inequality: organizations need to test who benefits, who bears risk, and whether human accountability remains meaningful as AI use expands (United Nations, 2015).

Conclusion

Co-Intelligence is a strong practitioner oriented starting point for structured experimentation with AI. Its most durable contribution lies in the Jagged Frontier and the

Centaur, Cyborg, and task allocation concepts, which make human AI collaboration concrete enough to test in real settings. Its most important unresolved issue is the boundary between useful and potentially harmful anthropomorphism. For practitioners, that tension creates a clear agenda for action: identify where human like framing is being used, compare alternative interaction designs, measure verification and decision quality, and revise the workflow when trust or delegation exceeds what the evidence justifies. Action research can therefore extend Mollick's framework by testing how organizations gain the benefits of AI while preserving human judgment, accountability, and agency.

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. (2022). *Falling asleep at the wheel: Human/AI collaboration in a field experiment on HR recruiters*. Laboratory for Innovation Science, Harvard Business School.
<https://static1.squarespace.com/static/604b23e38c22a96e9c78879e/t/62d5d9448d061f7327e8a7e7/1658181956291/Falling+Asleep+at+the+Wheel+-+Fabrizio+DellAcqua.pdf>
- Dell'Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K. C., 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>
- Goddard, K., Roudsari, A., & Wyatt, J. C. (2014). Automation bias: Empirical results assessing influencing factors. *International Journal of Medical Informatics*, 83(5), 368-375. <https://doi.org/10.1016/j.ijmedinf.2014.01.001>
- Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio/Penguin.
- United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development* (A/RES/70/1). <https://sdgs.un.org/2030agenda>
- Xie, Y., Zhao, S., Zhou, P., Lu, L., Liang, C., & Jiang, L. (2024). Estimating the impact of "humanizing" AI assistants. *International Journal of Human-Computer Interaction*, 40(24), 8876-8889. <https://doi.org/10.1080/10447318.2023.2291614>