Journal on Baltic Security, 2025; 11(2): 157-171

Book Review

Open Access
https://doi.org/10.57767/jobs_2025_007

Eden Elizabeth Armstrong*

Book Review: Ethan Mollick's Co-Intelligence: Living and Working with AI (2024)

Abstract: This review assesses Ethan Mollick's Co-Intelligence: Living and Working with AI (2024), which conceptualises artificial intelligence as a collaborative "co-intelligence" rather than a mere tool or autonomous agent. While recognising the book's clarity, accessibility, and transparency of both AI's potential and limitations, the review adopts a cautious, albeit critical, perspective in an effort to balance against Mollick's overarching optimism. It foregrounds concerns related to safety, equity, sustainability, and risk. The book's discussion of education is identified as a key strength, particularly its emphasis on AI literacy and human oversight. Considered through the lens of Professional Military Education (PME), Mollick's advocacy for experimentation is shown to sit uneasily with risk-averse, high-stakes environments. Nonetheless, the review concludes that Co-Intelligence offers a timely, pragmatic, and valuable contribution to debates on human-AI interaction.

Keywords: Artificial Intelligence, AI Integration, AI Ethics, Professional Military Education, book review.

* **Corresponding author:** Eden Elizabeth Armstrong, e-mail eden.armstrong@baltdefcol.org, Baltic Defence College

'Never trust anything that can think for itself if you can't see where it keeps its brain' – Mr. Weasley to his daughter Ginny in *Harry Potter and the Chamber of Secrets* (Rowling, 1998, p.329).

For those unfamiliar with the Harry Potter universe, Arthur Weasley is a wizard who is deeply fascinated by the non-magical world. Yet, in an eyebrow-raising turn of events, the man who canonically did not understand the function of a rubber duck, offers us a sage warning as we traverse our rapidly evolving more-science less-fiction reality.

I believe Mr. Weasley would find Artificial Intelligence (AI) mesmerising, yet terrifying; intriguing, yet suspicious…and likely full of Dark Magic. While AI cannot (yet?) "think for itself", Ethan Mollick's *Co-Intelligence: Living and Working with AI* (2024) enters an expanding body of literature that seeks to understand the sweeping implications of large language models (LLM) and generative artificial intelligence for not only the present, but also the future.

The book positions itself neither as a technical manual nor as a speculative treatise on artificial intelligence, but as a pragmatic guide to what Mollick characterises as a new era of 'co-intelligence', in which humans and AI systems jointly produce knowledge, decisions, and creative projects. The book's central claim – that generative AI represents an entirely new form of cognitive *partner* – is persuasive as it supports those interested and enthusiastic about AI integration, while also places dedicated emphasis on human recognition of AI's shortcomings, advocates for careful and ethical implementation, and validates core concerns for AI-sceptics.

In an effort for transparency, I find myself mostly in the latter camp. Not wholly. I am in no way sceptical that AI is transformational – it is. Or rather, it can be. I completely recognise AI has the potential to

change the way in which we live every day. However, at this specific moment in time, it is still only *potential*.

We are already seeing serious widespread ramifications stemming from improper development and implementation (see Marr, 2023; Myers, 2025; Zewe, 2024); negative learning outcomes that rival, and have the potential to be much more common than, the positive outcomes highlighted in Mollick's work (see Lin, 2024; Kilmova and Pikhart, 2025; Vieriu and Petra, 2025; Vilcarino and Langreo, 2025); not to mention the still developing environmental impacts (see United Nations Environment Programme, 2025; Griesser, 2025; Nutt, 2025; Wells, 2025); the unquantifiable psychological impacts of continued desocialisation (see Mollick, 2024; Brittian, 2025; Parham and Muller, 2025); and we are clearly lacking in the mental and emotional health space, as in the most devastating of circumstances, individuals have already lost their lives after their AI "companions" encouraged violence and suicide (see Forlini, 2025; Parham and Muller, 2025; Tabachnick, 2025; Schoene and Canca, 2025).

Ethical considerations and implementation issues are featured throughout many chapters (Mollick, 2024, pp. 27-62, 123-158). When read as an entire piece, a takeaway is that people's individual fears of replacement and oversight in addition to governmental and institutional lag are elements feeding AI scepticism. A minor challenge I offer to the book, is that AI-sceptics are *not* merely sceptics because they fear replacement or oversight. Not entirely anyway. As a contributing factor? Sure, some people may have an aversion to AI due to concerns surrounding replacement, oversight, or struggling to adapt to a new technology. However, as a reader in that AI-cautious camp, this notion comes off as somewhat missing the point. From a practical point of

view, many of us, AI-enthusiast and AI-sceptic alike, already operate under the assumption that we are replaceable in our workplaces. While Mollick highlights numerous areas of AI-integration issues, the overarching optimism of the piece diminishes valid concerns surrounding safety, security, sustainability, and equity in this transitional and transformational – if not revolutionary – time.

I can acknowledge that this minimisation could have been done as to not derail the book's premise of emphasising overarching potential. To his credit, Mollick does, repeatedly and strongly, emphasise elements of weakness within the AI-ethos we find ourselves in. Hopefully his optimism, married with my practical – albeit existential – concerns, can offer a middle ground between the potential of both promise and peril; a tightrope that we are walking if we like it or not.

Before You Assume I Hated the Book…

I did not hate it. In fact, I deeply appreciate it and its perspective that runs perpendicular to my own. Mollick may not have changed my view nor transformed me into an AI-enthusiast, however, reading *Co-Intelligence* did allow me *numerous* opportunities to challenge my own preconceived notions as well as perspectives. That alone, especially for those who are also AI-cautious, is reason enough to read and critically engage with the work. The book stands out for its clarity, pedagogical ambition, and obvious effort at balanced engagement of both strengthens and weaknesses of AI. At its core, the book argues that contemporary AI systems should be understood as *imperfect collaborators* whose value emerges through interaction and engagement rather than as tools that simply execute predefined tasks.

Mollick contends that effective use of AI requires new forms of literacy, experimentation, and institutional adaptation (Mollick, 2024, pp. 46-62). This framing underpins the book's practical orientation as readers are

encouraged to actively work with AI systems, probe their limits, and develop norms for responsible deployment in a wide variety of everyday contexts. While this orientation may be overarchingly optimistic, Mollick does repeatedly emphasise uncertainty, fallibility, and the need for human oversight, dispelling any potential for a straightforward techno-utopian reading.

A recurring virtue of the work is this candour regarding uncertainty. Mollick directly details that current AI systems are unstable, opaque, and subject to rapid change. Rather than presenting fixed prescriptions, he advocates for continuous learning and institutional flexibility. This epistemic humility distinguishes the book from more declarative or speculative accounts of AI futures. At the same time, the focus on adaptability may inadvertently normalise a state of perpetual disruption, placing the burden of adjustment on individuals and organisations rather than on developers and regulators.

Of all the genuinely numerous strengths of this book, the most notable lies in its accessibility. Mollick avoids extended technical exposition, yet he succeeds in conveying key concepts – such as probabilistic text generation, source and fact hallucination, and emergent behaviour – in a manner that is both accurate and intelligible to non-specialists. For an academic audience, this may appear elementary at times; however, the restraint shown in the engagement and more casual writing style is not only appropriate, but in my opinion, excellent, given the book's ambitions. This very accessibility is what allows us the space to attempt to apply the principles outlined in the work to Professional Military Education (PME) as well as the military and defence sectors more broadly.

Implications for Professional Military Education

While, the book's largely civilian organisational frame limits its direct applicability to military and defence sector institutions, within the context of PME, *Co-Intelligence* can be read as both a practical resource and a conceptual provocation. Mollick's emphasis on human-AI collaboration aligns with longstanding PME concerns regarding decision-making under uncertainty, staff work, and the integration of new technologies into command-and-control processes.

Mollick's treatment of education is particularly compelling (2024, pp. 159-177). Rather than framing generative AI as an existential threat to learning, he argues that AI exposes longstanding weaknesses in assessment practices and instructional design. As an Associate Professor at the University of Pennsylvania, Mollick integrated AI into his courses, with what appears to be smashing success (2024, pp. V-XV; 166-168). Recognising students are going to use AI, regardless of its institutional permissibility, Mollick seems to have cracked the code to AI integration in a traditional classroom setting, highlighting the success that his students have achieved in the last few years. The book brilliantly weaves Mollick's personal and professional anecdotes throughout the work, demonstrating to readers the ways in which his views on working with AI have been shaped while also allowing us to see the development of LLMs at the same time.

His acknowledgment that assignments, assessments, and the classroom work must adapt to emphasise process, critique, and metacognition is well aligned with existing educational research, and his willingness to concede that genuine risks – such as overreliance, deskilling, and erosion of trust – exist adds credibility to his position. Even readers who remain unconvinced that AI will ultimately enhance learning may find this chapter valuable for its concrete and empirically grounded

recommendations. These sorts of tangible recommendations offer PME educators a useful lens for rethinking curricula, particularly in relation to staff processes, wargaming, and the teaching of critical thinking in AI-augmented environments. By grounding the necessity of AI literacy, scepticism, and active supervision into the future of education, *Co-Intelligence* complements rather than displaces existing PME objectives and effectively offers a bridge between the civilian and military education systems.

However, I would argue, the biggest fault to this book, when it comes to a PME application, is Mollick's optimism regarding experimentation and iterative learning as it sits uneasily with risk-averse cultures and high-stakes consequences that are at the foundation of military operations – and therefore professional military education. While his insistence that AI systems function best as fallible partners rather than authoritative decision-makers resonates with doctrinal principles that privilege human judgment, responsibility, and ethical accountability in the use of force; there is not (currently) an answer for risk mitigation as we, more or less, fumble our way through this rapid technological expansion.

For educators overarchingly, we can take solace in knowing that if students make mistakes in our courses, there are unlikely to be disastrous real-world consequences – that is the point of making mistakes in an educational setting. However, for *PME educators*, it is our pinnacle responsibility, to ensure that our students not only learn, but *master*, the concepts we teach them before they leave the safety of our institutional walls and deploy the techniques in scenarios that *will* have real-world consequences.

Conclusion

I maintain deeply held concerns that AI is not well enough understood for its integration, much less rushed and haphazard integration, into core systems like healthcare, education, or the military and defence sectors. However, an equally strong core tenant of mine is that we *must* start with education. For PME, we recognise that militaries can only be as strong, resilient, and quality as their individual and collective training, making AI and technological education in PME institutions of the utmost importance as we navigate the rapidly changing technological landscape.

Despite these limitations, it is important to recognise Mollick's success on his own terms. *Co-Intelligence* does not aim to be a comprehensive theory of AI or a manifesto for technological reform.[1] Instead, it functions as an interpretive and practical intervention at a moment of widespread confusion and polarised debate. Mollick's writing is measured, evidence-informed, and free of both alarmism and hype. For scholars who are critical of prevailing narratives surrounding AI, the book offers a useful foil: a thoughtful, well-argued articulation of semi-cautious more-so optimism grounded in everyday practice.

Co-Intelligence: Living and Working with AI is a timely contribution to contemporary discussions of human-AI interaction. The book excels in clarifying what is genuinely new about generative AI and in offering a framework for responsible engagement. Even readers who fundamentally disagree with Mollick's outlook will find value in his careful reasoning, empirical awareness, and refusal to reduce complex phenomena to simplistic conclusions. As such, the book deserves

[1] See AI Statement for an oddity with this sentence.

serious consideration within academic debates on technology, work, and knowledge in the age of artificial intelligence.

AI Declaration Statement:

This AI Declaration Statement is both more information than likely necessary as well as particularly casual in tone in an effort for clarity and transparency. When I took on this book review, I had no intention of using AI. However, after reading just the first few pages of Mollick's work, doing things how I have always done them was only* going to result in a wasted opportunity for experimentation, open-mindedness, thoughtful engagement with Mollick's ideas, and ultimately, for me to learn something new – even if it pushed me outside of my comfort zone.

I specifically used ChatGPT, Perplexity, and NotebookLM. In each of these platforms, I asked for summaries of Mollick's work. Mollick details an interaction with ChatGPT regarding a New York Times article (2024, pp.76-79) so I gave a similar interaction a go. To similar results, the AI platforms were unable to provide quality summaries. The platforms tried, but most of their summaries were quite nonsensical for someone who had actually read the book – and at times were out-right inaccurate with hallucinated elements.

Next, I gave my 10 most prominent thoughts and asked the AI platforms to offer suggested outlines for the book review. The results across the platforms were okay, though would have resulted in an uninteresting book review – in my opinion. To my surprise ChatGPT offered the most standard book review outline. I had an (unfounded) assumption that NotebookLM would offer the most ordinary outline. I

did not use any of the suggested outlines but rather opted for a free writing style for my initial drafts.

Once I had what I, more or less, wanted to say written out, I asked all three platforms to offer suggestions on tone, diction, and syntax to overarchingly mediocre results. The suggestions were fine, a bit bland, and sometimes missed linguistic nuance and therefore the suggestions offered were inappropriate contextually. However, when I asked for suggestions on examples (specifically *not* sources; though some hallucinated links were brought forth anyway); many of the ideas offered did not make sense from the book's perspective; from a broad education perspective; much less a PME perspective – all of which were discarded.

Thus far in the process, I had an internal sense that the ChatGPT recommendations were the ones I preferred. I have no data to prove that this was the case or that they were substantively better than any other suggestion from any of the other platforms. From this point I used exclusively ChatGPT.

As I was implementing some of the suggested changes, my writing felt like it was getting increasingly messy, repetitive, and disjointed. While this happens in the writing process anyway, it was uncomfortable as this is the first piece I have written with an imperfect AI collaborator.

This is where it went off the metaphorical rails: I specifically asked ChatGPT to: 1) restructure and reorganise the book review; 2) to highlight, *not change*, just to point out, places of repetition for me to edit; and 3) to point out places that lacked continuity and read as disjointed. ChatGPT absolutely gave this a go, but (after what is countless tries) I did not think that the review was getting better. If I am blunt, I hated every suggestion more than the last.

And then, to my very, very, very deep irritation – and honestly chagrin – I found places where ChatGPT – unprompted – *changed* wording, sentence structure, some content, and for some reason inserted 'em dashes' everywhere without indicating it made changes. When questioned on this, ChatGPT apologised, but continued to make these sorts of changes despite being explicitly told not to. I have left a sentence that ChatGPT changed on its own to highlight it in this AI declaration:

- Original Sentence: '*Co-Intelligence* does not aim to be, nor claim to be, a comprehensive theory of AI; nor a playbook nor guide for technological reform'.
- ChatGPT Sentence: '*Co-Intelligence* does not aim to be a comprehensive theory of AI or a manifesto for technological reform'.

While ChatGPT's version may stand as an accurate statement; to me the – unprompted and unwanted – changes made carry a different theoretical intention and connotation as well as a different linguistic meaning to what was originally written.

During the editing process, Guest Editor William "Bill" Combes' Claude AI assistant suggested I professionalise the tone and offered some recommendations. Here is one of the suggestions; the only editing done in the following paragraph was to format to comply with British English:

'The original framing ('nor claim to be... nor a playbook') emphasises Mollick's explicit rhetorical choices through parallel negation. ChatGPT's revision ('or a manifesto') simplifies the syntax but loses the accumulated emphasis. More significantly, "manifesto" carries political

connotations absent from 'playbook nor guide for technological reform'. These subtle shifts demonstrate how AI optimisation toward clarity can inadvertently alter argumentative tone' – Claude AI Assistant.

While I happen to overarchingly agree with what I interpret as Claude's intentions; 1) there is no reference to 'Mollick's explicit rhetorical choices'; 2) the 'parallel negation' was my stylistic choice; and 3) I am not entirely convinced of the value-added by expanded linguistic analysis. I had already mentioned that the linguistic changes made had an impact on meaning and intention. Did we *need* to go into specifics?

Where mistakes were made: Unfortunately, I did not notice the unrequested changes made by ChatGPT until approximately 5 or 6 reorganisations in. I was not looking for language and content changes because I did not request ChatGPT make language and content changes. I requested ChatGPT point out areas for improvement. I had specifically asked for recommendations. I thought I was specific enough in my request for restructuring and reorganisation of content present. Apparently, I was not.

I want to stress that this comes down to entirely user-error oversights in 2 ways: 1) not being specific enough in my prompts and 2) not thoroughly enough checking AI's work as we went along. Rookie mistakes on both fronts.

The tone of the book review I intended to give was genuine appreciation for Mollick's work, despite my harsh-at-times critiques. I wanted to lift up the successes of the book while offering additional elements for all of us to seriously consider as we navigate the rapidly changing environment alongside AI. I now harbour concern that the sorts of changes made by ChatGPT before I noticed may have minimised my intended meaning. I have gone back and tried to reassemble my original review from what I had put into the AI

platforms. I have read and reread this book review in an attempt to capture all the places that were changed; however, it would be arrogant of me to assume complete success in this endeavour.

The last element of unease for me is that now – because *I* made mistakes in detailed continuity verification – I cannot be absolutely, unequivocally sure of what was AI-produced and what was Eden-produced. Maybe the concern feels bigger than it actually is. Maybe my AI-cautious stance is clouding my judgement. Regardless of my internal battle, I still no longer feel as though I am the sole author. While this may complicate authenticity and authorship, an element excellently brought forward in Mollick's work and by Claude AI; it would be disingenuous of me to minimise this complication and to downplay my mistakes for the sake of "academic standardisation and language professionalisation". Regardless of where we draw the line on human-AI collaboration as per authorship, I, without question nor hesitation, take full accountability for what is written.

Bibliography

Brittain, B. (2025) OpenAI Sued for Allegedly Enabling Murder-Suicide. Reuters. [online] 11 Dec. Available at: https://www.reuters.com/legal/government/openai-sued-allegedly-enabling-murder-suicide-2025-12-11/ (Accessed 18 December 2025).

Forlini, E. (2025) OpenAI Sued by 7 Families for Allegedly Encouraging Suicide, Harmful Delusions. [online] PCMAG. Available at: https://www.pcmag.com/news/openai-sued-by-7-families-for-allegedly-encouraging-suicide-harmful-delusions (Accessed 18 December 2025).

Griesser, K. (2025) Study Reveals How Much Energy AI Uses to Answer Your Questions. [online] CNN. Available at: https://www.cnn.com/2025/06/22/climate/ai-prompt-carbon-emissions-environment-wellness (Accessed 20 November 2025).

Klimova, B. and Pikhart, M. (2025) Exploring the Effects of Artificial Intelligence on Student and Academic Well-Being in Higher Education: A Mini-Review. Frontiers in Psychology, [online] 16(16). doi: https://doi.org/10.3389/fpsyg.2025.1498132.

Lin, L. (2024) A Quarter of U.S. Teachers Say AI Tools Do More Harm than Good in K-12 Education. [online] Pew Research Center. Available at: https://www.pewresearch.org/short-reads/2024/05/15/a-quarter-of-u-s-teachers-say-ai-tools-do-more-harm-than-good-in-k-12-education/ (Accessed 15 December 2025).

Mollick, E. (2024) Co-Intelligence: Living and Working with AI. New York, New York: Penguin Random House.

Muller, R.T. and Parham, A. (2025) Is Artificial Intelligence Perpetuating Loneliness? [online] Psychology Today. Available at: https://www.psychologytoday.com/us/blog/talking-about-trauma/202509/is-artificial-intelligence-perpetuating-loneliness?msockid=357144056693617d1bc151c467c460ae (Accessed 19 December 2025).

Myers, A. (2025) Researchers Work to Squash 'Fantastic Bugs' Hidden in AI Benchmarks. [online] Standford Report. Available at: https://news.stanford.edu/stories/2025/12/ai-benchmarks-flaws-models-bugs-research (Accessed 16 December 2025).

Ngila, F. (2025) Studies Show AI Has Been Encouraging Suicide and Fueling Mental Illnesses. [online] Impact Newswire. Available at: https://impactnews-wire.com/studies-show-ai-has-been-encouraging-suicide-and-fueling-mental-illnesses/ (Accessed 18 December 2025).

Nutt, D. (2025) 'Roadmap' Shows the Environmental Impact of AI Data Center Boom. [online] Cornell Chronicle. Available at: https://news.cornell.edu/stories/2025/11/roadmap-shows-environmental-impact-ai-data-center-boom (Accessed 21 November 2025).

Rowling, J.K. (1998) Harry Potter and the Chamber of Secrets. London, England: Bloomsbury Publishing.

Tabachnick, C. (2025) Parents of Teens Who Died by Suicide After AI Chatbot Interactions Testify in Congress. [online] CBS News. Available at: https://www.cbsnews.com/news/ai-chatbots-teens-suicide-parents-testify-congress/ (Accessed 19 December 2025).

United Nations Environment Programme. (2024) AI Has an Environmental Problem. Here's What the World Can Do About That. [online] United Nations Environment Programme. Available at: https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about (Accessed 20 November 2025).

Vieriu, A.M. and Petrea, G. (2025) The Impact of Artificial Intelligence (AI) on Students' Academic Development. Education Sciences, [online] 15(3), p.343. doi:https://doi.org/10.3390/educsci15030343.

Vilcarino, J. and Langreo, L. (2025) Rising Use of AI in Schools Comes With Big Downsides for Students. [online] Education Week. Available at: https://www.edweek.org/technology/rising-use-of-ai-in-schools-comes-with-big-downsides-for-students/2025/10 (Accessed 2 December 2025).

Wells, C. (2025) Growing Use of AI Brings an Environmental Impact. [online] AP News. Available at: https://apnews.com/article/ai-data-center-climate-impact-environment-c6218681ffdbad5bf427b47347fddcb9 (Accessed 22 November 2025).