



© Ipopba-AdobeStock

## B. Business Impact

# Friend and Foe: Cognitive, Emotional and Moral Consequences of Working with AI

ESCP Impact Paper No.2023-09-EN

Chi HOANG, Hsin-Hsuan MEG LEE & Daniela LUP  
ESCP Business School



ESCP Impact Paper No.2023-09-EN

## **Friend and Foe: Cognitive, Emotional and Moral Consequences of Working with AI**

Chi Hoang\*  
Hsin-Hsuan Meg Lee\*\*  
Daniela Lup\*\*\*

ESCP Business School

### **Abstract**

As the power of artificial intelligence (AI) continues to grow, businesses face a crucial decision: embrace its potential benefits or risk being left behind. But with great power comes great responsibility, and this paper explores the cognitive, emotional, and moral consequences of working with AI. We argue that as AI at work has the potential to reshape the very essence of human decision-making, organizations have the responsibility to proactively mitigate any adverse effects and thoughtfully weigh the pros and cons of AI deployment.

Keywords: artificial intelligence, cognitive assimilation, emotional AI, moral agency

\*Assistant Professor of Marketing, ESCP Business School

\*\*Associate Professor of Marketing, ESCP Business School

\*\*\*Professor of Human Relations and Organizational Behavior, ESCP Business School

ESCP Impact Papers are in draft form. This paper is circulated for the purposes of comment and discussion only. Hence, it does not preclude simultaneous or subsequent publication elsewhere. ESCP Impact Papers are not refereed. The form and content of papers are the responsibility of individual authors. ESCP Business School does not bear any responsibility for views expressed in the articles. Copyright for the paper is held by the individual authors.

## **Friend and Foe: Cognitive, Emotional and Moral Consequences of Working with AI**

As artificial intelligence (AI) becomes ubiquitous in our personal and professional lives, managers must grapple with its impact on their organizations and employees. AI is now widely used for various tasks such as screening and hiring, diagnosing diseases, and providing customer services. A survey conducted by McKinsey shows that nearly two-thirds of the surveyed companies were planning to increase their heavy investments in AI in the workplace in the coming years (McKinsey, 2021). Noting this significant trend, we wonder how working with AI will impact human workers cognitively, emotionally, and morally. In this paper, we discuss three different scenarios concerning the psychological and behavioural consequences of working with AI.

### **The cognitive scenario: AI cognitive assimilation**

AI is designed to mimic human intelligence. As technology evolves, the latest versions of AI demonstrate high computer processing capacity that exceeds that of the human brain in many areas. A crucial aspect of human intelligence is its ability to adapt, learn new concepts from examples, and leverage prior knowledge to enable novel solutions (Barbey, 2018). This capability is referred to as cognitive flexibility, and this, AI is reportedly lacking. Take AlphaGo as an example. After AlphaGo defeated a human world champion in Go, its parent company failed to improve the AI further because AlphaGo could not adapt or learn new moves by itself (Silver & Hassabis, 2017). Given AI's cognitive limitation, what will happen when human workers regularly interact and collaborate with AI, a cognitively inflexible version of their mind?

It is important to note that exposures to others, human or nonhuman, tend to trigger changes in one's own cognition and behaviors. Humans tend to unconsciously assimilate into others to facilitate more effective interactions with others. An example of this is the process of cultural assimilation through which immigrants who live with people from a culture other than their own start adopting some of the norms and behaviors of their new culture. Assimilation also happens at the organizational level, whereby individuals learn and adjust their beliefs and behaviors to the normative organizational practices demonstrated by their co-workers (Kramer & Miller, 1999). Humans also assimilate to technologies. For instance, millennials' behaviors are being rewired to assimilate to software structures, rapidly searching, browsing, assessing quality, and synthesizing vast quantities of information (Anderson & Rainie, 2012). Moreover, internet users, after searching for information online, show an increase in self-assessed knowledge, evidence of them assimilating the characteristics of the internet into their self-concept (Fisher, et al., 2015). As such, working with AI is likely to also alter the human cognitive structure, especially on the cognitive flexibility dimension that distinguishes humans from AI. Thus, we argue that human workers may become cognitively less flexible as a consequence of working with AI.

Importantly, AI-induced cognitive inflexibility may exert downstream consequences on human workers' behaviors. Cognitively flexible individuals tend to exhibit lower commitment behavior, as documented in the religious domain whereby religious disbelief and disengagement are related to higher cognitive flexibility (Zmigrod et al., 2019). As such, AI-induced cognitive inflexibility may possibly increase human workers' commitment behaviors, such as higher persistence in difficult tasks but also more resistance to organizational changes and innovations. Cognitive flexibility is also an important antecedent of creativity (Nijstad et al., 2010); thus working with AI may inhibit individuals'

creativity at work. Furthermore, cognitive flexibility is a driver of immoral behaviors, as cognitively flexible individuals can better justify their immoral actions (Gino & Ariely, 2012). As such, being inflexible may offer a moral benefit (i.e., reinforcement of employees' moral behavior). Overall, these examples show that cognitively assimilating to AI has both positive and negative behavioural outcomes which need to be purposefully taken into consideration when organizations plan to deploy AI in the workplace.

### **The emotional scenario: Emotional AI paradox**

Emotions play a pivotal role in how individuals respond to any technology, including AI. Although AI is neither human nor as nuanced as humans are, it is certainly able to elicit or mitigate emotional reactions from people. Recent studies have found that interaction with AI elicits both negative effects such as anxiety and disappointment, but also positive emotions such as happiness and excitement (Shank et al., 2019). In some cases, people respond to AI with more intense emotions than to humans. Given that emotions guide actions, it is important to understand how human decision-makers are emotionally influenced by AI. In general, the more human-like an AI system is, the more emotional the response it generates. But how human-like does it have to be?

AI anthropomorphism refers to the practice of treating AI as human, not just by designing AI interfaces with human-like facial features but also by seeing AI as having personalities, mental states or feelings. Emotional intelligence describes the capability to recognize, understand and regulate one's emotions and those of others. While current AI systems are incapable of emotions and feelings, some are trained to detect and process human emotions. Emotional AI, developed from affective computing, has made significant progress in recent years, and this allows AI to exhibit features that can be perceived as emotional intelligence. AI can also identify emotions through facial expression recognition, voice and language processing, and other physiological signals and, although it cannot feel emotions, the algorithmic design can make an emotional expression that resembles human patterns. For example, an AI-powered chatbot, Wysa, can offer supportive messages based on the emotions it detects and is programmed to show patience, empathy, and comfort. Many other similar tools have been used for human service and care.

The growing use of emotional AI in tracking and analyzing human emotions, behaviors, and attitudes is often promoted as a way to improve well-being. However, the current use of emotional AI has raised many unresolved issues. For example, the problematic accuracy in emotion detection of these tools can create false trust and distrust in the system and in the organizations and practices that use AI in this manner, with consequences harmful for both individuals who take decisions and those whom their decisions affect. We don't suggest that the way forward is to remove emotional AI from use, but rather that we need to understand it better.

The advancement of AI can confront our sense of agency in thinking about what makes us humans. Affective agency, the ability to have emotions, is central to human identity. This is also what we tend to expect from our counterparts in interactions. Paradoxically, the emotional effects of human-AI interactions are found to be diminished (when positive), remedied, or improved (when negative) if the individuals are made aware that they interact with AI. The awareness leads to a perceived lack of affective agency with AI and consequently leads to resistance to emotional engagement. Along with asking if AI can engage emotionally, the question of "what" makes meaningful emotional interactions should be addressed first as this is a structural problem that the development of emotional AI needs to consider. Emotional involvement and commitment from both parties in an exchange serve as the foundation of balanced social interactions. However, when only one

is emotionally invested or knows the other well, as in all social exchanges with AI, this one-sided interaction could feel like involuntary parasocial interaction with many known negative consequences to well-being. This issue can be signified by the observations of why humans do not always feel social emotions, such as empathy (Liu-Thompkins, Okazaki & Li, 2022) and jealousy when interacting with AI.

### **The moral scenario: Working with AI and moral agency**

The moral agency represents the agency exercised by individuals in pursuit of actions aligned with their moral values. In practice, it manifests in a dual form: either as proactive engagement in such actions or as refusal to engage in actions and behaviours seen as immoral (Bandura, 2002). The moral agency does not arise in a vacuum; it is socially situated, which means that the characteristics of one's social context affect the specific ways in which individuals engage in or disengage from questioning the moral implications of their pursuits. Given that AI is likely to continue altering the working context, and notwithstanding the harm that decisions devoid of moral judgment could inflict on employees, organisations, and society, we propose that, before deploying AI at work, it is important to understand whether and how such changes can affect employees' exercise of moral agency.

Scholars of ethics have identified several context characteristics, the presence of which are necessary for individuals to develop and exercise moral agency (e.g., Bandura, 2002; MacIntyre, 2000). Below we discussed two characteristics that we see as crucial for establishing a climate conducive to exercising moral agency when dealing with AI.

First, the working context should allow the decision makers to exercise self-reflection, to 'debate with themselves' whether the decision suggested by AI responds to moral values that they hold high as individuals, not only as organizational players. Take the example of a manager who, based on AI evaluations embedded in the HRM system, would make employees redundant. Or, one that is free from immediate harm, in which the AI suggests substantial promotion to several employees. Questions of moral agency appear regardless of the type of outcome. In each of these situations, to exercise moral agency would mean, at the minimum, questioning the criteria that underlie the recommended decision, thinking whether they would have reached the same decisions, why and based on what values (e.g., care, fairness etc.). Yet, the veil of objectivity and efficiency that shrouds decisions made by AI is not conducive to such self-reflection. In fact, reflecting on AI decisions would contradict its very role of objective and efficient solution to human decision failures. Thus, contexts in which AI is framed as objective and effective (most of the use cases of AI) are unlikely to stimulate the exercise of self-reflection and rather encourage delegation of agency to the AI.

Second, contexts conducive to human moral agency provide space and scope for deliberation about the shared purpose and values that guide decisions. They require the presence of "reliable others" (MacIntyre, 2000) and the "psychological safety" (Edmondson, 1999) that allow for challenging the logic of objectivity assumed in AI processes. It is with regard to this dimension that current research offers limited advice. We know for instance that group decisions in the presence of experts are likely to end up endorsing the expert opinion, but we have no research evidence about the extent to which human decision-makers defer to the 'AI expert'. We suggest however that the more organisations frame AI tools as objective evaluators, the less likely it is for group members to bring to the fore moral judgment – which are viewed as subjective – and thus the more likely that the AI's decision will remain unchallenged. Furthermore, the more AI objectivity is linked to goals of efficiency and profit, the less likely it is for individuals to develop and exercise moral agency.

## Conclusion

Our analysis suggests three take-aways for organizations that want to make the most of the AI revolution without crowding out unique humanity dimensions.

1. Ensure a balance between working with AI and working with humans to mitigate negative consequences of AI cognitive assimilation, such as diminished cognitive flexibility or creativity.
2. Pay attention to unintended negative emotions that can be elicited by emotional AI, as well as the unresolved issues surrounding its usages, such as inaccuracy in emotion detection and one-sided emotional engagement.
3. Strive to create working contexts that avoid moral delegation to AI and maximize individuals' ability to exercise and further develop moral agency.

## References

- Anderson, J., & Rainie, L. (2012). Millennials will benefit and suffer due to their hyperconnected lives. Washington DC, Pew Research Center, 18.
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, 31(2), 101-119.
- Barbey, A. K. (2018). Network neuroscience theory of human intelligence. *Trends in Cognitive Sciences*, 22(1), 8-20.
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350-383.
- Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the Internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General*, 144(3), 674-687.
- Gino, F., & Ariely, D. (2012). The dark side of creativity: original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, 102(3), 445-459.
- Kramer, M. W., & Miller, V. D. (1999). A response to criticisms of organizational socialization research: In support of contemporary conceptualizations of organizational assimilation. *Communication Monographs*, 66(4), 358-367.
- Liu-Thompkins, Y., Okazaki, S., & Li, H. (2022). Artificial empathy in marketing interactions: Bridging the human-AI gap in affective and social customer experience. *Journal of the Academy of Marketing Science*, 50(6), 1198-1218.
- MacIntyre, A. (2000). "After Virtue: A Study in Moral Theory (2nd ed.)". London: Duckworth
- McKinsey. (2021). The state of AI in 2021. Retrieved from <https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2021>
- Nijstad, B. A., De Dreu, C. K., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European review of social psychology*, 21(1), 34-77.

Shank, D. B., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior*, 98, 256-266.

Silver, D., & Hassabis, D. (2017). AlphaGo Zero: Starting from scratch. Deepmind. Retrieved from <https://www.deepmind.com/blog/alphago-zero-starting-from-scratch>

Zmigrod, L., Rentfrow, P. J., Zmigrod, S., & Robbins, T. W. (2019). Cognitive flexibility and religious disbelief. *Psychological Research*, 83(8), 1749-1759.