Editorial: The ethics and behavioral economics of human-AI interactions

Marina Chugunova[1] and Nicola Lacetera[2]*

[1]Department for Innovation and Entrepreneurship Research, Max Planck Institute for Innovation and Competition, Munich, Germany, [2]Department of Management, University of Toronto, Toronto, ON, Canada

Editorial on the Research Topic
The ethics and behavioral economics of human-AI interactions

This Research Topic is a collection of studies of the rapidly evolving landscape of human-AI interactions, a subject that has attracted increasing attention as artificial intelligence becomes intertwined with our daily lives. The proliferation of AI applications has not only amplified the convenience and efficiency of our routines, but has also raised critical ethical questions and triggered changes in human behavior.

Although some patterns already documented for interactions with previous generations of technologies are likely to extend to the current wave of AI, some of its features warrant specific examination. In particular, the ability of AI systems to continuously learn from new data and experiences means that they can evolve over time and even in real time, offering contextually relevant interactions and providing information that are tailored to the individual user's needs. On the one hand, this changes the performance expectations of the user, but on the other hand, it makes the outcomes less predictable, and the process more opaque, than in the interaction with older versions of automated agents. In essence, the special quality of AI lies in its mimicry of human learning processes and its adaptability to the user. This feature opens a space for strategic interactions on both sides: Human users may adjust their behavior to generate desirable outcomes, for example to affect individualized pricing; AI agents might adjust their responses to increase engagement, for instance by offering the information that the user is more likely to like, thus potentially fostering and amplifying biases, creating echo chambers, and spreading disinformation.

These peculiarities raise questions and concerns not for a distant future; they are immediate and pressing as AI technologies become more capable and widespread. How, for example, is cooperation achieved when humans interact with "artificial agents"? What is different or similar as compared to human-human interactions? Do people display similar or different behavioral tendencies and biases [other regarding preferences, time preferences, risk attitudes, (over)confidence, etc.] when working with artificial agents as compared to humans? What are people's attitudes toward the use of intelligent machines for certain tasks or functions? What moral concerns does this raise? What are the reasons for any potential opposition to the reliance on AI-operated machines for certain tasks?

Behavioral economics offers approaches to understand the nuanced ways in which interacting with AI affects human behavior. The papers in this Research Topic highlight the breadth of questions to be addressed: from the role of human personality traits for the hybrid

interactions, to reliance on technology, intergroup dynamics and immoral behavior. The findings from these studies as well as from many ongoing research efforts remind us that this interaction is not a simple case of mechanical replacement but a fundamental transformation of the decision landscape. AI's influence on human behavior is intricate and often counterintuitive. The presence of AI alters the context in which decisions are made, the information that is available, and the strategies that are employed.

Various foundational methods in behavioral economics, such as laboratory and field experiments, have been employed to provide causal evidence on the topic. These methods effectively abstract from and control for potential confounding factors that might be challenging or unfeasible to keep constant using observational data. In addition, new tools—such as field-in-the-lab experiments—allows investigating real-world interactions in a controlled environment. Taking stock of existing evidence and theoretical contributions, moreover, conceptual analyses can offer unique insights from a number of the regularities documented in previous studies.

The interaction with AI is dynamic and evolving due to the rapid pace of technological change. Although the exact sizes of the estimated effects might be context-specific and may change from one generation of a technology to another, we can and should study underlying behavioral regularities that are persistent and shape the general framework of the interaction with technology.

The overarching narrative is clear: the rise of AI is not just a technological or economic phenomenon, but a behavioral one. The research presented here is united by a common goal: to navigate the ethical and economic implications of our deepening relationship with AI. The insights gleaned from these and many other studies to come can help pave the way for a future where AI and human behavior co-evolve in a manner that is beneficial and, above all, human-centric.

Author contributions

MC: Writing – original draft, Writing – review & editing. NL: Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note