

Abstract:

Human–AI interaction is as much about how people think as about what AI does. Expectations and self-monitoring shape how users engage with intelligent systems and interpret their own performance. Our research shows that merely believing in AI support can enhance both confidence and task performance, even when no real assistance is provided— a placebo effect of AI. This optimism proves remarkably resilient: even when AI is described as unreliable or stress-inducing, people continue to expect improvement and improve their task performance, demonstrating the robustness of the effect. More recent work reveals a new paradox. While AI reliably enhances cognitive performance, it simultaneously impairs users' insight into their own task accuracy.

People tend to overestimate their competence, and those with greater AI literacy are often the most overconfident. Taken together, these findings suggest that the next challenge for human-centred Human–AI interaction is to design for human metacognition, to help users remain aware of their interaction with AI.