

People hold smart AI assistants responsible for outcomes, study finds

August 30 2023



Credit: *iScience* (2023). DOI: 10.1016/j.isci.2023.107494

Even when humans see AI-based assistants purely as tools, they ascribe partial responsibility for decisions to them, as a new study shows.

Future AI-based systems may navigate autonomous vehicles through traffic with no human input. Research has shown that people judge such futuristic AI systems to be just as responsible as humans when they make autonomous traffic decisions. However, real-life AI assistants are far removed from this kind of autonomy. They provide human users with supportive information such as navigation and driving aids.

So, who is responsible in these real-life cases when something goes right or wrong? The human user? Or the AI assistant? A team led by Louis Longin from the Chair of Philosophy of Mind has now investigated how people assess responsibility in these cases. Their findings are published in *iScience*.

"We all have smart assistants in our pockets," says Longin. "Yet a lot of the experimental evidence we have on responsibility gaps focuses on robots or autonomous vehicles where AI is literally in the driver's seat, deciding for us. Investigating cases where we are still the ones making the final decision, but use AI more like a sophisticated instrument, is essential."

A philosopher specialized in the interaction between humans and AI, Longin, working in collaboration with his colleague Dr. Bahador Bahrami and Prof. Ophelia Deroy, Chair of Philosophy of Mind, investigated how 940 participants judged a human driver using either a smart AI-powered verbal assistant, a smart AI-powered tactile assistant, or a non-AI navigation instrument. Participants also indicated whether

they saw the navigation aid as responsible, and to which degree it was a tool.

Ambivalent status of smart assistants

The results reveal an ambivalence: Participants strongly asserted that smart assistants were just tools, yet they saw them as partly responsible for the success or failures of the human drivers who consulted them. No such division of responsibility occurred for the non-AI powered instrument.

No less surprising for the authors was that the smart assistants were also considered more responsible for positive rather than negative [outcomes](#).

"People might apply different moral standards for praise and blame. When a crash is averted and no harm ensues, standards are relaxed, making it easier for people to assign credit than blame to non-human systems," suggests Dr. Bahrami, who is an expert on collective responsibility.

Role of language is not relevant

In the study, the authors found no difference between smart assistants that used language and those that alarmed their users by a tactile vibration of the wheel.

"The two provided the same information in this case, 'Hey, careful, something ahead,' but of course, ChatGPT in practice gives much more information," says Ophelia Deroy, whose research examines our conflicting attitudes toward [artificial intelligence](#) as a form of animist beliefs. In relation to the additional information provided by novel language-based AI systems like ChatGPT, Deroy adds, "The richer the

interaction, the easier it is to anthropomorphize."

"In sum, our findings support the idea that AI assistants are seen as something more than mere recommendation tools but remain nonetheless far from [human](#) standards," says Longin.

The authors believe that the findings of the new study will have a far-reaching impact on the design and social discourse around AI assistants: "Organizations that develop and release smart [assistants](#) should think about how social and moral norms are affected," Longin concludes.

More information: Louis Longin et al, Intelligence brings responsibility—Even smart AI assistants are held responsible, *iScience* (2023). [DOI: 10.1016/j.isci.2023.107494](https://doi.org/10.1016/j.isci.2023.107494)

Provided by Ludwig Maximilian University of Munich

Citation: People hold smart AI assistants responsible for outcomes, study finds (2023, August 30) retrieved 10 May 2024 from <https://techxplore.com/news/2023-08-people-smart-ai-responsible-outcomes.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.
