

Using cues and actions to help people get along with artificial intelligence

March 31 2020, by Matt Swayne



Learning how people interact with AI is an important step in integrating machines into society in beneficial ways. Credit: Pixabay

Learning how people interact with artificial intelligence-enabled machines—and using that knowledge to improve people's trust in

AI—may help us live in harmony with the ever-increasing number of robots, chatbots and other smart machines in our midst, according to a Penn State researcher.

In a paper published in the current issue of the *Journal of Computer-Mediated Communication*, S. Shyam Sundar, James P. Jimirro Professor of Media Effects in Donald P. Bellisario College of Communications and co-director of the Media Effects Research Laboratory, has proposed a way, or framework, to study AI, that may help researchers better investigate how people interact with artificial intelligence, or Human-AI Interaction (HAI).

"This is an attempt to systematically look at all the ways [artificial intelligence](#) could be influencing users psychologically, especially in terms of trust," said Sundar, who is also an affiliate of Penn State's Institute for Computational and Data Sciences (ICDS). "Hopefully, the theoretical model advanced in this paper will give researchers a framework, as well as a vocabulary, for studying the social psychological effects of AI."

The framework identifies two paths—cues and actions—that AI developers can focus on to gain trust and improve user experience, said Sundar. Cues are signals that can trigger a range of mental and emotional responses from people.

"The cue route is based on superficial indicators of how the AI looks or what it apparently does," he explained.

Sundar added that there are several cues that affect whether users trust AI. The cues can be as obvious as the use of human-like features, such as a [human face](#) that some robots have, or a human-like voice that virtual assistants like Siri and Alexa use.

Other cues can be more subtle, such as a statement on the interface explaining how the device works, as in when Netflix explains why it is recommending a certain movie to viewers.

But, each of these cues can trigger distinct mental shortcuts or heuristics, according to Sundar.

"When an AI is identified to the user as a machine rather than human, as often happens in modern-day chatbots, it triggers the 'machine heuristic,' or a mental shortcut that leads us to automatically apply all the stereotypes we hold about machines," said Sundar. "We might think machines are accurate and precise, but we also might think of computers and machines as cold and unyielding." These stereotypes in turn dictate how much we trust the AI system.

Sundar suggested that autopilot systems in airplanes are one example of how over-trust in AI can lead to negative repercussions. Pilots may trust so implicitly in the autopilot system that they relax their guard and are not prepared for sudden changes in the plane's performance or malfunctions that would require their intervention. He cites this kind of 'automation bias' as an indication of our deep trust in machine performance.

On the other hand, AI can also trigger negative biases for some people.

"The opposite of automation bias would be algorithm aversion," said Sundar. "There are people who just have an aversion because, perhaps in the past, they were burned by an algorithm and now deeply mistrust AI. They were probably fooled by 'deepfakes' which are fabricated videos created using AI technology, or they got the wrong product recommendation from an e-commerce site, or felt their privacy was invaded by AI snooping into their prior searches and purchases."

Sundar advised developers to pay particular attention to the cues they may be offering users.

"If you provide clear cues on the interface, you can help shape how the users respond, but if you don't provide good cues, you will let the user's prior experience and folk theories, or naive notions, about algorithms to take over," Sundar said.

In addition to providing cues, the AI's ability to interact with people can also fashion [user experience](#), according to Sundar. He calls this the "action route."

"The action route is really about collaboration," said Sundar. "AIs should actually engage and work with us. Most of the new AI tools—the smart speakers, robots and chat bots—are highly interactive. In this case, it's not just visible cues about how they look and what they say, but about how they interact with you."

In both actions and cues, Sundar suggests developers maintain the correct balance. For example, a cue that does not transparently tell the user that AI is at work in the device might trigger negative cues, but if the cue provides too much detail, people may try to corrupt—or "game"—the interaction with the AI. "Cueing the right amount of transparency on the interface is therefore quite important," he said.

"If your smart speaker asks you too many questions, or interacts with you too much, that could be a problem, too," said Sundar. "People want collaboration. But they also want to minimize their costs. If the AI is constantly asking you questions, then the whole point of AI, namely convenience, is gone."

Sundar said he expects the framework of cues and actions to guide researchers as they test these two paths to AI trust. This will generate

evidence to inform how developers and designers create AI-powered tools and technology for people.

AI technology is evolving so fast that many critics are pushing to outright ban certain applications. Sundar said that giving researchers the time to thoroughly investigate and understand how humans interact with technology is a necessary step to help society tap the benefits of the devices, while minimizing the possible negative implications.

"We will make mistakes," said Sundar. "From the [printing press](#) to the internet, new media technologies have led to negative consequences, but they have also led to many more benefits. There is no question that certain manifestations of AI will frustrate us, but at some point, we will have to co-exist with AI and bring them into our lives."

More information: S Shyam Sundar. Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAI), *Journal of Computer-Mediated Communication* (2019). [DOI: 10.1093/jcmc/zmz026](https://doi.org/10.1093/jcmc/zmz026)

Provided by Pennsylvania State University

Citation: Using cues and actions to help people get along with artificial intelligence (2020, March 31) retrieved 19 April 2024 from <https://techxplore.com/news/2020-03-cues-actions-people-artificial-intelligence.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.