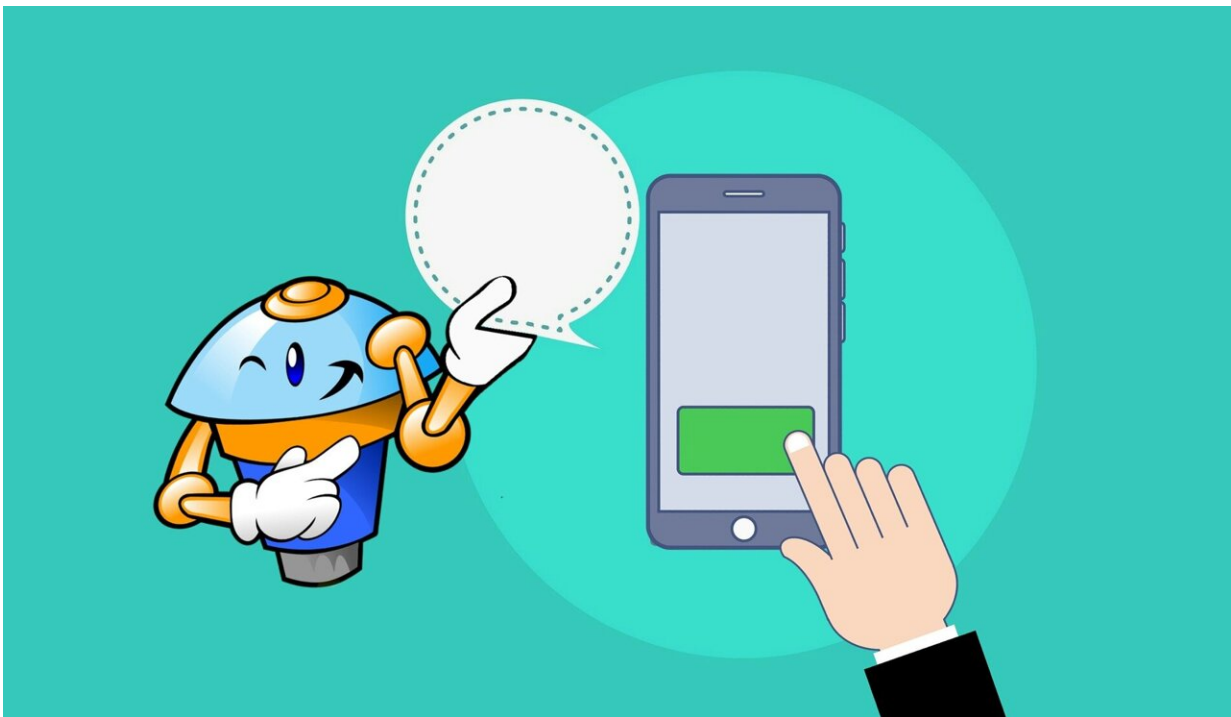# What influences trust when conversing with chatbots?

January 9 2025, by Noemi Kern

Credit: Pixabay/CC0 Public Domain

Whether on your bank's website or your telephone provider's help line, interactions between humans and chatbots have become part of our daily lives. But do we trust them? And what factors influence our trust? Researchers at the University of Basel recently examined these questions.

"Hello ChatGPT, can you help me?"—"Of course, how can I help you?" Exchanges between users and chatbots, which have their basis in artificial intelligence (AI), quickly seem like conversations with another person.

Dr. Fanny Lalot and Anna-Marie Betram from the Faculty of Psychology at the University of Basel wanted to know how much people trust AI chatbots and what this trust depends on. They focused on text-based systems—that is, platforms like ChatGPT rather than voice assistants such as Siri or Alexa.

Test subjects were exposed to examples of interactions between users and a chatbot called Conversea that was imagined specifically for the study. They then imagined they would interact with Conversea themselves. The results are published in the *Journal of Experimental Psychology: General*.

The chatbot as an independent entity

Our level of trust in other people depends on a variety of factors: our own personality, the other person's behavior and the specific situation all play a role. "Impressions from childhood influence how much we are able to trust others, but a certain openness is also needed in order to want to trust," explains social psychologist Lalot. Characteristics that promote trust include integrity, competence and benevolence.

The new study shows that what applies to relationships between humans also applies to AI systems. Competence and integrity in particular are important criteria that lead humans to perceive an AI chatbot as reliable. Benevolence, on the other hand, is less important, as long as the other two dimensions are present.

"Our study demonstrates that the participants attribute these

characteristics to the AI directly, not just to the company behind it. They do think of AI as if it was an independent entity," according to Lalot.

Additionally, there are differences between personalized and impersonal chatbots. If a chatbot addresses us by name and makes reference to previous conversations, for example, the study participants assessed it as especially benevolent and competent.

"They anthropomorphize the personalized chatbot. This does increase willingness to use the tool and share personal information with it," according to Lalot. However, the test subjects did not attribute significantly more integrity to the personalized chatbot and overall trust was not significantly higher than in the impersonal chatbot.

Integrity is more important than benevolence

According to the study, integrity is a more important factor for trust than benevolence. For this reason, it is important to develop the technology to prioritize integrity above all else. Designers should also take into account the fact that personalized AI is perceived as more benevolent, competent and human in order to ensure proper use of the tools. Other research has demonstrated that lonely, vulnerable people in particular run the risk of becoming dependent on AI-based friendship apps.

"Our study makes no statements about whether it is good or bad to trust a chatbot," Lalot emphasizes. She sees the AI chatbot as a tool that we have to learn to navigate, much like the opportunities and risks of social media.

However, there are some recommendations that can be derived from their results. "We project more onto AI systems than is actually there," says Lalot. This makes it even more important that AI systems be reliable. A chatbot should neither lie to us nor endorse everything we say

unconditionally.

If an AI chatbot is too uncritical and simply agrees with everything a user says, it fails to provide reality checks and runs the risk of creating an echo chamber that, in the worst case, can isolate people from their social environment. "A [human] friend would hopefully intervene at some point if someone developed ideas that are too crazy or immoral," Lalot says.

Betrayed by AI?

In human relationships, broken trust can have serious consequences for future interactions. Might this also be the case with chatbots? "That is an exciting question. Further research would be needed to answer it," says Dr. Lalot. "I can certainly imagine that someone might feel betrayed if advice from an AI chatbot has negative consequences."

There need to be laws that hold the developers responsible. For example, an AI platform could show how it comes to a conclusion by openly revealing the sources it used, and it could say when it doesn't know something rather than inventing an answer.

More information: Fanny Lalot et al, When the bot walks the talk: Investigating the foundations of trust in an artificial intelligence (AI) chatbot., *Journal of Experimental Psychology: General* (2024). DOI: 10.1037/xge0001696

Provided by University of Basel