

ELLA GLIKSON

# EMOTIONAL OVER-TRUST IN AI TECHNOLOGY

Our trust in AI technology is based on an evaluation of how we feel about it and how it performs. However, when we cannot evaluate its technological performance, this emotional-based trust can easily become a problematic over-trust.

We need a minimal level of trust to use any type of new technology. Some of this trust is based on rational thinking (e.g., the new technology's predicted reliability and usefulness), and some trust is grounded in emotion (e.g., linked to the extent we like the way the new technology is presented).

“...anthropomorphic features such as facial features, human-like voice, or physical form significantly increase liking and trust.

User Interface (UI) specialists use psychological principles to make technology easier to use and improve its attractiveness and likeability. One of the most popular ways they can make AI more likeable is to emphasise different types of anthropomorphic features. Empirical research consistently demonstrates that factors such as facial features, human-like voice, or physical form significantly increase liking and trust.

When these human-like features (such as AI's immediate responsive behaviours to user movement or words) also signal a technological ability to perform the required task, the two elements of trust (liking and rationality) are aligned. However, what happens when the presence of human-like features overshadow the rational evaluation of the technological ability, and why is this important?

## EMOTIONAL VS RATIONAL THINKING

Human-like cues lead to high expectations about AI's technological performance. Research shows that the more technology is presented as a living organism, the more we like it and believe in its capabilities and moral values. For instance, giving an automated car a human name can increase our liking and trust, and lead to assumptions about the car's high performance and reliability. However, the level of technological attractiveness is based on the employment of psychological principles (e.g., similarity to a living thing or to a specific user) and has little to do with its algorithmic functions.

“ AI's human-like behaviours likely impact our emotions more profoundly than our rational thinking.

AI's human-like behaviours likely impact our emotions more profoundly than our rational thinking. In several studies, researchers found that people tend to trust anthropomorphic robots, even when their low-performance ability was evident. Although these studies were performed in labs, where the actual implication of robotic performance is questionable, they raise an important query regarding the relative power of the emotional basis of trust in technology. The more complex the outcomes of algorithmic performance, the more difficult it is to correctly evaluate reliability, and thus the role emotions play in the evaluation become more significant.

## PREVENTING OVER-TRUST

The disassociation between technology's likeability and its actual reliability and performance can be highly problematic, resulting in over-trust. Over-trust relates to a situation in which high trust in unreliable technology would lead to misuse, which may cause a breach of safety or other undesirable outcomes. Based on people's tendency to resist change, research tends to focus on ways to improve trust and facilitate the use of new technologies. Therefore, a lot of effort is made to understand how to improve the likeability of bots and robots and make them more integrated into organisations and everyday life.

Although this effort can result in an eagerness to use new technologies, there is a growing need to better understand how to balance the positive emotions evoked by technology's external features and the need for a rational evaluation of technology's reliability and performance.

Aiming to find new ways in which our (manipulated) emotional reactions will not lead to over-trust in technology that is biased, erroneous, or just not yet ready to perfectly perform the task at hand, we need to put more effort in demonstrating this phenomenon in lab and on-line experiments, as well as communicating the possible dangers to those with responsibility for purchasing new technology, and potentially also those with responsibility for regulating its use.

*Ella Glikson is an assistant professor at the Graduate School of Business Administration in Bar Ilan University.*