

Association for Information Systems

AIS Electronic Library (AISeL)

WHICEB 2023 Proceedings

Wuhan International Conference on e-Business

Summer 5-28-2023

When Intelligent Voice Assistants fail: The mitigating role of verbal and vocal warmth

Bo Huang

University of Nottingham Ningbo China, 199 Taikang East Road, Ningbo 315100, China,
bo.huang@nottingham.edu.cn

Sylvain Sénécal

HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, Quebec, Canada H3T 2A7

Follow this and additional works at: <https://aisel.aisnet.org/whiceb2023>

Recommended Citation

Huang, Bo and Sénécal, Sylvain, "When Intelligent Voice Assistants fail: The mitigating role of verbal and vocal warmth" (2023). *WHICEB 2023 Proceedings* . 44.

<https://aisel.aisnet.org/whiceb2023/44>

This material is brought to you by the Wuhan International Conference on e-Business at AIS Electronic Library (AISeL). It has been accepted for inclusion in WHICEB 2023 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

When Intelligent Voice Assistants fail: The mitigating role of verbal and vocal warmth

Bo Huang¹ and Sylvain Sénécal²

¹ University of Nottingham Ningbo China, 199 Taikang East Road, Ningbo 315100, China

² HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, Quebec, Canada H3T 2A7

bo.huang@nottingham.edu.cn (Bo Huang, corresponding author)

1. INTRODUCTION AND RESEARCH QUESTIONS

Intelligent voice assistants (IVAs), such as Alexa and Siri, are speech-enabled integrated artificial intelligence (AI) technologies that allow for voice-based conversational interaction between consumers and the interface. The adoption of these voice assistants is on the rise both in the home (personal) context (e.g., Alexa and Google Home) and the service (commercial) context (e.g., voice-based assistants and chatbots). However, despite the exciting growth and increasing popularity of IVAs, extant literature on human-computer interaction (HCI) suggests that the key technologies behind them --- namely Natural Language Processing (NLP) and Automatic Speech Recognition (ASR) face greater challenges and limitations because of the complexity of human language. For example, IVAs suffer from the known limits of deep learning algorithms, meaning that they can only work in the specific domains where data training was provided for. When facing unfamiliar requests, they are likely to fail or act erratically. Therefore, since these negative incidents are bound to happen and could lead to discouragement or even total abandon, how can we better handle interaction failures, especially in a consumer context? This research intends to answer the question by drawing from both Computers As Social Actors (CASA) research paradigm ^{[1][2]} and Stereotype Content Model ^{[3][4]}.

2. THEORY AND RESEARCH FRAMEWORK

The Computer As Social Actors (CASA) research paradigm ^{[1][2]} suggests that the human-computer relationship is fundamentally social: people apply the same social heuristics used for human interactions to computers. We argue that if consumers indeed interact with IVAs as if they were interacting with humans, two universal dimensions of social cognition should be relevant in this context: warmth and competence^{[3][4]}. Research suggests that people generally have a positive attitude toward those who are perceived as warmer. Recent studies on HCI have shown that such a positive effect of warmth extends to interaction with non-human entities such as robots. For example, prior research has found that people have similar emotional and behavioral reactions to robots as they have to humans: they are more likely to help the robots when they are perceived as less competent but warmer ^[5].

Taken together, in line with the previous research on warmth, we posit that it could also affect how consumers react when the interaction outcome is rather negative. A failed interaction can be viewed as a sign of low competence since IVAs cannot comprehend, process, or fulfill what is requested by the consumer. Therefore, a voice assistant that is perceived as warm but performs poorly can be regarded as an individual with high warmth but low competence – it has good intentions to carry out the tasks but lacks the ability to do them. Compared to a voice assistant in which both warmth and competence are lacking (e.g., seen as both cold and incompetent), the former should elicit less negative emotional reactions from consumers.

3. RESULTS AND MAJOR FINDINGS

We conducted two experiments. In the first experiment, results from both physiological and psychological measures suggest that warmth has a positive impact in mitigating the negative emotional reactions from an interaction failure. Specifically, we find that a higher perceived warmth leads to a higher level of emotional arousal, as well as a reduced frustration felt by the participants. In the second study, we replicate the effects in a service failure context, and results

additionally suggested that the benefits of warmth go beyond consumers' emotional responses, as they report higher re-patronage intention despite a service failure.

4. CONTRIBUTIONS

Theoretically, the current research makes a primary contribution to our understanding of service failure involving IVAs. There is a growing stream of empirical research on consumer interaction with AI-enabled technologies such as algorithms, IVAs, and service robots, yet little attention has been paid to consumer-AI service encounters, particularly when those services fail. We, therefore, contribute to the emerging literature on how to reduce the negative consequences of service failure involving an "AI employee", by showing that creating an impression of warmth can improve overall emotional reactions and increase re-patronage intention. This directly extends past research that found that consumers expect greater warmth from humanoid service robots^[6]. Second, while prior research on HRI suggests that various factors affect the judgment of warmth for non-human entities, the effect of different vocal features is less clear. We fill this gap by comparing two speech delivery styles (conversational vs dynamic) identified in human speech. We found that a voice assistant receives the highest warmth rating when it is verbally warm and vocally dynamic. Third, extant literature calls for more diverse methods and utilizing emerging tools to fully understand consumer emotional experience^[7]. The current research answers this call by incorporating psychophysiological measures (i.e., EDA) as one of the emotional measures in our study. This non-intrusive neuroscience tool allows us to tap into the consumer's emotional state during the very moment of interaction with a voice assistant, as opposed to subjective feelings of emotion after interaction which past research mostly focused on.

REFERENCES

- [1] Nass, Clifford Ivar, and Scott Brave (2005), "Wired for speech: How voice activates and advances the human-computer relationship," Cambridge, MA: MIT press.
- [2] Nass, Clifford, and Youngme Moon (2008), "Machines and mindlessness: Social responses to computers," *Journal of Social Issues*, 56(1), 81-103.
- [3] Fiske, Susan T., Amy JC Cuddy, and Peter Glick (2007), "Universal dimensions of social cognition: Warmth and competence," *Trends in Cognitive Sciences*, 11(2), 77-83.
- [4] Fiske, Susan T., Amy JC Cuddy, Peter Glick, and Jun Xu (2002), "A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition," *Journal of Personality and Social Psychology*, 82(6), 878-902.
- [5] Mieczkowski, Hannah, Sunny Xun Liu, Jeffrey Hancock, and Byron Reeves (2019), "Helping not hurting: Applying the stereotype content model and BIAS map to social robotics," In 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 222-229.
- [6] Choi, S., Mattila, A. S., & Bolton, L. E. (2021). To err is human (-oid): how do consumers react to robot service failure and recovery?. *Journal of Service Research*, 24(3), 354-371.
- [7] Caruelle, Delphine, Anders Gustafsson, Poja Shams, and Line Lervik-Olsen (2019), "The use of electrodermal activity (EDA) measurement to understand consumer emotions—a literature review and a call for action," *Journal of Business Research*, 104, 146-160.