How Can Real-World Feedback and Priming Affect Trust in Simulated Autonomous Agents?

Hiroshi Furuya*
University of Central Florida

Jasmine Joyce DeGuzman
University of Central Florida

Zubin Choudhary
University of Central Florida

Matt Gottsacker
University of Central Florida

Gerd Bruder
University of Central Florida

Gregory F. Welch
University of Central Florida

Abstract

Autonomous agents present important novel capabilities for a wide swath of applications like simulating interactions between humans and agents. Simulating these interactions in VR has become an important tool for evaluating the effects of agents on human behavior and performance, including human-agent trust. This position paper presents research opportunities in the use of real-world multi-modal feedback and real-world priming experiences may have on the validity of trust measurements taken from simulated human-agent interactions in VR. In addition, it presents a hypothetical experiment addressing research questions related to this topic.

Index Terms: Autonomous agents, trust, presence, virtual reality.

1 Introduction

Autonomous agents are leading advancement in a wide range of disciplines and domains. By simulating these agents in virtual reality (VR), researchers can emulate and measure the effects of agent designs and behaviors. One of the most important factors for successful human-agent interaction is trust, defined as the "attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" [10]. Trust is especially important in team-based, mission-oriented activities [10] where any miscalibration of the agent's actual performance capabilities and the user's trust in the agent results in poor overall performance [4, 12].

Although crucial in the success of human-agent teams, trust is considered a latent, difficult-to-measure property [8]. Furthermore, the number of definitions on trust that agree on the inclusion of a component of risk, vulnerability, or uncertainty has led to calls for risk being explicitly included to ensure the validity of the trust measurement [18]. If a user does not believe that trusting the agent can potentially aid in their aims, the concept of trust in the agent loses its meaning. Thus, it is necessary to achieve participant "buy-in" to enable meaningful measurements of trust. Even with the large, growing body of literature addressing the impacts of VR system characteristics on social interactions with agents [3, 6], there is still little that addresses human-agent trust and what aspects of VR experiences impact the validity of trust evaluations.

In order to even attempt to measure an elusive variable like trust, it is crucial that users perceive palpable risk or uncertainty such that it elicits a "response-as-if-real (RAIR)" [15]. Simulating autonomous agents in VR with highly realistic environments makes it possible to evoke these desired responses. Additionally, highly realistic environments provide experimenters with high levels of ecological validity for their studies [9].

*e-mail: {hiroshi.furuya, jasdeg, zubin.choudhary, mattg, bruder, welch}@ucf.edu

It is essential for VR experiences that study human-agent interactions to be realistic enough for users to perceive actual risk and uncertainty in order to induce a meaningful trust RAIR. This work considers multi-modal feedback and user interaction priming techniques as potential trust RAIR influences in VR simulations of human-agent interactions. Leveraging multimodal feedback, such as external haptic or olfactory devices, and real-world experiences to prime users for interactions with the agent prior to a VR experience ostensibly improves presence [1, 2]. By improving the level of presence, it also improves the likelihood that a user would find the interaction scenario credible, thus providing a trust RAIR. This work then uses these factors to pose a hypothetical experiment that evaluates their influence on evaluating trust.

2 Multi-modal Feedback and Presence

Presence is classically defined by the two components place illusion (PI) –"the qualia of having a sensation of being in a real place" [16]– and plausibility illusion (Psi) – "the illusion that the scenario being depicted as actually occurring" [15]. Thus, increased presence contributes to more authentic RAIR. Prior work has also demonstrated that increased presence improves the capacity of a VR experience to elicit emotional responses [11]. This is important to note because trust, along with risk perception and uncertainty, is largely an affective process [10].

It has been shown that multimodal feedback improves presence, particularly plausibility [1, 5]. Feedback can be supplied to users in VR by using external devices. For example, haptic feedback can be provided by a fan that blows air on users in windy virtual environments [7] and olfactory feedback can be provided by scented objects in environments where smells are expected.

3 Real-World Priming

Previous work suggests that contextual priming, such as the prior reading of an article related to the VR experience, can improve presence in VR [2]. It is possible that real-world experiences could also be used as a type of contextual primer as psychology has demonstrated that presence has links to memory outcomes [17]. For example, imagine if a user were to interact with a real robot in a real environment and experienced a simulation of the same robot in a virtual environment that mirrored the real world. Would this interaction elicit a more authentic RAIR than if the virtual environment did not match the one of the real world experience?

4 Hypothetical Experiment

We pose a hypothetical experiment that evaluates how the trust between a human and an autonomous agent, such as an unmanned aerial vehicle (UAV), is affected by multimodal environmental feedback and various priming techniques. It seeks to address the following research questions: **R1**: Does improving plausibility through multi-modal environmental feedback improve the validity of simulated human-agent trust measurements? **R2**: What effect does real-world priming have on the validity of simulated human-agent trust measurements?

The experiment consists of three independent variables: *Priming: (3)* Real-world experience, textual prompt, none (baseline); *Agent reliability: (3)* Low, medium, high; and *Feedback: (2)* Audio-only, multi-modal.

Those in the real-world priming condition experience a small UAV flying near them such that they can hear the spinning propellers and even feel the wind that it creates. Meanwhile, the textual prompt condition has participants read a description of the UAV, its physical properties and how it might operate. The control group receives no priming material. In the audio-only conditions, participants would only hear the sound of the UAV flying in the air, while in the multi-modal conditions, haptic feedback and olfactory feedback would be added as well. The multi-modal feedback is based on elements placed in the environment and the properties of the UAV. A blowing fan would simulate the winds of the UAV propellers providing haptic feedback. Olfactory feedback can be implemented by placing an object in the virtual environment, such as a bouquet of flowers, and placing a flowery scent diffuser in the real world near the participant. All of the additional apparatus for this experiment can be seen in Figure 1.

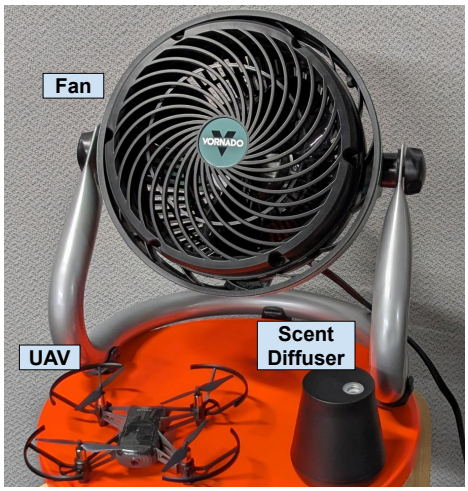

Figure 1: Clockwise from top: A fan used to simulate wind, a scent diffuser, and a miniature UAV

This experiment uses a mixed design where participants would be divided into 3 groups by Priming condition. Every group experiences each agent reliability and feedback condition in a 3×2 factorial design. During the experiment, participants would first receive priming based on their group. Then, participants would perform a trust task with a simulated UAV in VR for each of the reliability-feedback conditions. This could be a trust game [13], an observation of a UAV completing a task, or even a vignette [14] supported by an encounter with the simulated UAV similar to what the participants in the real-world priming condition would experience. Agent reliability would be manipulated in the context of this task. In addition to task-specific measures, such as performance in a trust game, we would utilize self-report measures to assess trust, perceived risk and uncertainty, and presence, among other factors.

Acknowledgments

This material includes work supported in part by the National Science Foundation under Award Number 2235066 (Dr. Han-Wei Shen, IIS); the Office of Naval Research under Award Numbers N00014-21-1-2578 and N00014-21-1-2882 (Dr. Peter Squire, Code 34); and the AdventHealth Endowed Chair in Healthcare Simulation (Prof. Welch).

References

[1] I. Bergstrom, S. Azevedo, P. Papiotis, N. Saldanha, and M. Slater. The Plausibility of a String Quartet Performance in VR. *IEEE Trans. on Vis. and Comp. Graphics*, 23(4):1352–1359, Apr. 2017. doi: 10.1109/TVCG.2017.2657138 1

[2] L. Cerda, A. Fauvarque, P. Graziani, and J. Del-Monte. Contextual priming to increase the sense of presence in virtual reality: exploratory study. *Virtual Reality*, 25(4):1105–1112, Dec. 2021. doi: 10.1007/s10055-021-00515-4 1

[3] A. Felnhofer, T. Knaust, L. Weiss, K. Goinska, A. Mayer, and O. D. Kothgassner. A Virtual Character's Agency Affects Social Responses in Immersive VR: A Systematic Rev. and Meta-Anal. *Intl. Jour. of HCI*, 40(16):4167–4182, Aug. 2024. doi: 10.1080/10447318.2023.2209979 1

[4] H. Furuya, L. Battistel, Z. D. Choudhary, M. Gottsacker, G. Bruder, and G. F. Welch. Difficulties in perceiving and understanding robot reliability changes in a sequential binary task. In *Proc. of 2024 ACM Symp. on Spatial User Interaction*, SUI '24, pp. 1–11. ACM, Trier, Germany, 2024. doi: 10.1145/3677386.3682083 1

[5] J. K. Gibbs, M. Gillies, and X. Pan. A comp. of the effects of haptic and visual feedback on presence in vr. *Intl. Jour. of Human-Computer Studies*, 157:102717, Jan. 2022. doi: 10.1016/j.ijhcs.2021.102717 1

[6] G. Jacucci, A. Bellucci, I. Ahmed, V. Harjunen, M. Spape, and N. Ravaja. Haptics in social interaction with agents and avatars in vr: A systematic rev. *Virtual Reality*, 28(4):170, Nov. 2024. doi: 10.1007/s10055-024-01060-6 1

[7] K. Kim, R. Schubert, J. Hochreiter, G. Bruder, and G. Welch. Blowing in the wind: Increasing social presence with a virtual human via environmental airflow interaction in mixed reality. *Elsevier Computers and Graphics*, 83(October 2019):23–32, 2019. doi: 10.1016/j.cag.2019.06.006 1

[8] S. C. Kohn, E. J. de Visser, E. Wiese, Y.-C. Lee, and T. H. Shaw. Measurement of Trust in Automation: A Narrative Rev. and Ref. Guide. *Frontiers in Psych.*, 12, 2021. 1

[9] C. Kyrlitsias and D. Michael-Grigoriou. Social Interaction With Agents and Avatars in Immersive Virtual Environments: A Survey. *Frontiers in VR*, 2, Jan. 2022. doi: 10.3389/frvir.2021.786665 1

[10] J. D. Lee and K. A. See. Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 46(1):50–80, Mar. 2004. doi: 10.1518/hfes.46.1.50_30392 1

[11] M. Meehan, B. Insko, M. Whitton, and F. P. Brooks. Physiological measures of presence in stressful virtual environments. *ACM TOG*, 21(3):645–652, July 2002. doi: 10.1145/566654.566630 1

[12] R. Parasuraman and V. Riley. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, 39(2):230–253, June 1997. doi: 10.1518/001872097778543886 1

[13] N. Rahimzadagan, M. Vahs, I. Leite, and R. Stower. Drone Fail Me Now: How Drone Failures Affect Trust and Risk-Taking Decisions. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 862–866. ACM, Boulder CO USA, Mar. 2024. doi: 10.1145/3610978.3640609 2

[14] T. Sato, J. Inman, M. S. Politowicz, E. T. Chancey, and Y. Yamani. The Influence of Viability, Independence, and Self-Governance on Trust and Public Acceptance of Uncrewed Air Vehicle Operations. *Proc. of Human Factors and Ergonomics Society Annual Meeting*, 67(1):51–56, Sept. 2023. doi: 10.1177/21695067231194326 2

[15] M. Slater. Place ill. and plausibility can lead to realistic behaviour in immersive ves. *Phil. Trans. of the Royal Soc. B: Bio. Sci.*, 364(1535):3549–3557, Dec. 2009. doi: 10.1098/rstb.2009.0138 1

[16] M. Slater, D. Banakou, A. Beacco, J. Gallego, F. Macia-Varela, and R. Oliva. A Separate Reality: An Update on Place Ill. and Plaus. in VR. *Frontiers in VR*, 3, June 2022. doi: 10.3389/frvir.2022.914392 1

[17] S. A. Smith and N. W. Mulligan. Immersion, presence, and episodic memory in virtual reality environments. *Memory*, 29(8):983–1005, Sept. 2021. doi: 10.1080/09658211.2021.1953535 1

[18] N. Tenhundfeld, M. Demir, and E. de Visser. Assess. of Trust in Automation in the "Real World": Reqs. for New Trust in Automation Meas. Techniques for Use by Practitioners. *Jour. of Cog. Eng. and Dec. Making*, 16(2):101–118, June 2022. doi: 10.1177/

15553434221096261 1