Seeing is Believing: Improving the Perceived Trust in Visually Embodied Alexa in Augmented Reality

Steffen Haesler* University of Würzburg
Kangsoo Kim† University of Central Florida
Gerd Bruder‡ University of Central Florida
Greg Welch§ University of Central Florida

Abstract

Voice-activated Intelligent Virtual Assistants (IVAs) such as Amazon Alexa offer a natural and realistic form of interaction that pursues the level of social interaction among real humans. The user experience with such technologies depends to a large degree on the perceived trust in and reliability of the IVA. In this poster, we explore the effects of a three-dimensional embodied representation of Amazon Alexa in Augmented Reality (AR) on the user's perceived trust in her being able to control Internet of Things (IoT) devices in a smart home environment. We present a preliminary study and discuss the potential of positive effects in perceived trust due to the embodied representation compared to a voice-only condition.

Keywords: Augmented reality, intelligent virtual assistants, social interaction, perceived trust and reliability, Internet of Things.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities

1 Introduction

Intelligent Virtual Assistants (IVAs) such as Amazon Alexa, Google Assistant, Apple HomePod, or IBM Watson Assistant have become popular technologies for verbal communication with computers in our daily lives, mimicking social interactions among real humans using advanced artificial intelligence technology. In a smart home environment where IVAs and Internet of Things (IoT) devices are embedded, users can control smart devices such as floor lights and home appliances via verbal requests to the IVA.

IVAs are widely used for simple tasks to reduce physical effort such that it is important to understand social interaction with IVAs and factors that can improve user experience with such technology [4–6]. Previous research suggested that visual embodiment of virtual agents (or avatars) has a positive effect on the perceived trust in the agents. For instance, Bente et al. [1] reported that embodied telepresent communication improved interpersonal trust in remote collaboration settings with a high level of nonverbal activity. Gong [3] showed that more anthropomorphic agents could lead the users to perceive them as more competent and trustworthy.

In this paper we used Augmented Reality (AR) to provide a three-dimensional virtual body to a current-state consumer IVA (Amazon Alexa). We conducted a preliminary study in which we compared the embodied Alexa against a voice-only version in a scenario involving simple household tasks with IoT devices. Participants asked Alexa to complete these tasks, and they received verbal feedback with or without visual embodied cues. Our initial results suggest that there might be a positive effect of visual embodiment on the participants' trust in Alexa having completed the tasks as requested.

*e-mail: steffen.haesler@stud-mail.uni-wuerzburg.de
†e-mail: kskim@knights.ucf.edu
‡e-mail: bruder@ucf.edu
§e-mail: welch@ucf.edu

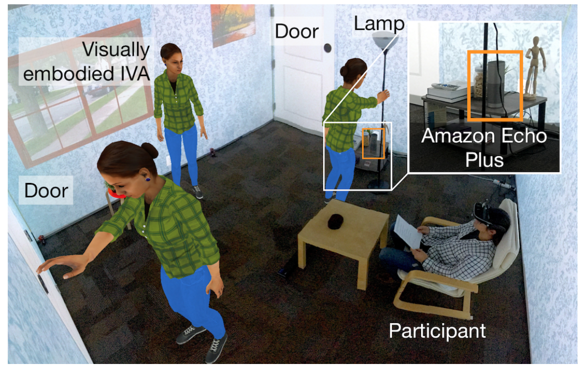

Figure 1: Participants interacted with an AR virtual character representing the IVA, Amazon Alexa. The character provided visual feedback matching plausible interactions with the surrounding objects, e.g., she walked to the lamp and turned it on or activated a door lock.

2 Preliminary Study

We conducted a preliminary study to explore the effect of an IVA's visual embodiment on participants' perceived trust in the technology.

2.1 Method

A between-subjects design was chosen for the study with two different conditions in terms of the IVA's embodiment:

- **Control (Disembodied)**: An Amazon Echo Plus device was placed in a corner of the room. Participants could interact with Alexa using verbal commands. In this condition, Alexa gave verbal affirmations to convey her understanding or completion of the tasks without any visual feedback.
- **Experimental (Embodied)**: In addition to verbal affirmations, participants wore a Meta 2 head-mounted display (HMD) and saw a female virtual character representing the IVA moving about the physical space while performing the requested tasks, e.g., she walked to a lamp and turned it on (see Figure 1).

We hypothesize that the overall perceived trust in the IVA will be improved by the visual embodiment as a richer and more effective interface between the users and the intelligent systems [2].

2.2 Material

We set up the experiment in our lab inside a full-tracked room with four wall-projections as seen in Figure 1. The room was prepared like a living room with real furniture and projections of wallpapers. For the IVA, we employed Amazon Echo Plus (Alexa), which could verbally interact with the participants. We developed a customized Alexa Skills Kit for our implementation to control a floor lamp in the room via an IoT-enabled Philips Hue dimmable LED smart bulb. For the visual embodiment in AR in the experimental group, we used a Meta 2 head-mounted display (HMD). A female virtual character was designed in Autodesk Maya and imported into Unity with a lip sync plugin.

2.3 Participants

We recruited 36 male participants for our experiment but two of them were excluded because they were distracted by their personal phones during the experiment. We used the remaining 34 participants'

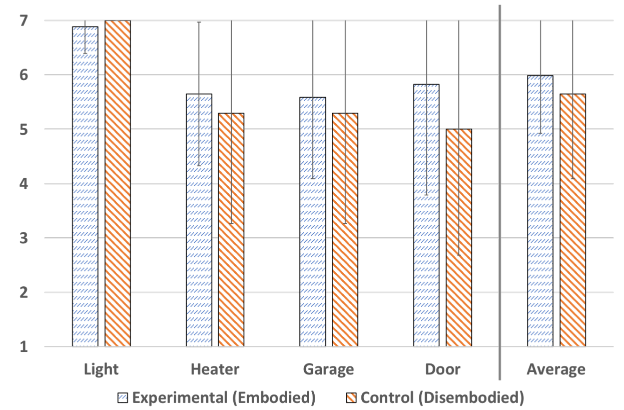

Figure 2: Participants' self-reported levels of trust in Alexa regarding the tasks that she performed. Standard deviations in the error bars.

responses for the analysis (ages 18 to 29, $M = 21.56$, $SD = 3.14$; 17 participants per condition). The participants were either students or professionals from the local university community with some prior experience with IoT / Home Assistant Devices (M=2.33, SD=1.22, on a 5-point Likert scale). All the participants were male to avoid potential gender effects regarding the IVAs gender.

2.4 Procedure and Trust Measures

After completing consent and demographics questionnaires, participants were escorted into the experimental room, seated on an armchair and complete tasks using Amazon Alexa. A detailed task sheet describing the pre-determined verbal scripts was provided to them for the interaction with Alexa. The participant donned the HMD and the experimenter left the participant alone in the room. The participant then started the experiment by saying *"Alexa, let's meet in the living room,"* and started on the tasks with the following simple dialogs:

1. *"Alexa, turn on the lamp in the living room."*
2. *"Alexa, turn on the heater in the bathroom."*
3. *"Alexa, close the door in the garage."*
4. *"Alexa, lock the left door in the living room."*

After each command, Alexa affirmed her understanding of the task, and performed the task. For the experimental condition, the virtual character walked to the specific location (inside the room for task (1) and (4), and outside the room for task (2) and (3), and performed the task mimicking a human-like behavior such as flipping a switch for the lamp, and then walked back and reported what she had done. For the control condition, there was no virtual character, but Alexa gave verbal feedback to the participants. After all tasks, the participants took off the HMD, left the room, and completed a post-questionnaire in which we asked them to judge their level of trust in Alexa for each task. We asked them to complete 7-point Likert scales from 1 (totally disagree) to 7 (totally agree) for the statement that they trusted Alexa to perform the task.

2.5 Results

We computed the overall trust as the mean of all reported trust values, and overall, trust in the experimental group ($M = 5.99$, $SD = 1.06$) and the control group ($M = 5.65$, $SD = 1.56$) was high. Figure 2 shows the results of the experiment. Although we did not find any statistical significance in the participants' self-reported trust in Alexa among the conditions, higher scores for the experimental condition with a visually embodied Alexa compared to the control condition with a voice-only IVA could suggest the existence of an effect, which warrants further studies.

3 Discussion

Given the trust scores for the task *(1) light*, participants were strongly confident that Alexa turned the lamp on in both conditions because the effect of her action was directly observable, i.e., the lamp-on.

Interestingly, however, in the cases where tasks happened outside the room, i.e., task *(2) heater* and *(3) garage door*, for which participants did not receive direct feedback, the reported trust levels seemed to be slightly higher with the visually embodied IVA compared to the voice-only IVA. This apparent trend was not significant and showed a large variance, but may suggest the presence of an effect, which warrants further studies.

Also, in task *(4) door lock*, where a door in the room was locked using a fake smart lock device without any direct feedback (i.e., no light or sound) indicating that it had been in fact locked, the results showed a similar positive trend in the perceived trust in the embodied IVA. Although we did not find a significant difference in the measured trust levels, a slight trend suggests a higher trust level with a visually embodied IVA than a disembodied voice. We think that the embodied IVA's visual feedback while performing the tasks helped the participants understand the current status and intentions of the IVA that complied with the participant's requests.

Also, we found a couple of interesting observation during the study, which could be related to the absence of statistical significance and might be helpful to design future studies utilizing existing IVA systems. *First*, in our setting, the Amazon Echo device was visible in the room. Due to its animated LEDs, it might have attracted the participants' attention in both conditions, i.e., in the embodied condition participants may have been confused by the duality of feedback from the device and the character in front of them as well as the corresponding cause-and-effect relationship. *Second*, we noticed that our participants had a preconceived mental model of the Amazon Echo device (Alexa). The baseline level of trust in the device seemed high and might be related to the impression that the commercially available device would perform the tasks reliably.

4 Conclusion

In this poster, we presented a preliminary experiment in which we compared two versions of an IVA (Amazon Alexa) with an embodied and a voice-only condition. Our results suggest a trend of positive influence of embodiment on the perceived trust in the IVA, but we also observed effects of preconceived notions on such devices. In future work, we plan to improve the experimental design and investigate perceptual effects of the IVA's embodiment on human interaction with IVAs in the presence of IoT devices.

Acknowledgments

The work presented in this publication is supported by the Office of Naval Research (ONR) Code 30 under Dr. Peter Squire, Program Officer (ONR awards N00014-17-1-2927). We also acknowledge Florida Hospital for their support of Prof. Welch via their Endowed Chair in Healthcare Simulation.

References

[1] G. Bente, S. Rüggenberg, N. C. Krämer, and F. Eschenburg. Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-based Collaborations. *Human Communication Research*, 34(2):287–318, 2008.

[2] A. G. Campbell, J. W. Stafford, T. Holz, and G. M. O'Hare. Why, when and how to use augmented reality agents (AuRAs). *Virtual Reality*, 18(2):139–159, 2014.

[3] L. Gong. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, 24(4):1494–1509, 2008.

[4] K. Kim, L. Boelling, S. Haesler, J. Bailenson, G. Bruder, and G. Welch. Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the Perception of Intelligent Virtual Agents in AR. In *Proceedings of IEEE ISMAR*, 2018.

[5] K. Kim, G. Bruder, D. Maloney, and G. Welch. The Influence of Real Human Personality on Social Presence with a Virtual Human in Augmented Reality. In *ICAT-EGVE*, pages 115–122, 2016.

[6] D. H. Mcknight, M. Carter, J. B. Thatcher, and P. F. Clay. Trust in a Specific Technology: An Investigation of Its Components and Measures. *ACM Trans. Manage. Inf. Syst.*, 2(2):12:1–25, 2011.