

Exploring Virtual Agents for Augmented Reality

Isaac Wang

Department of CISE
University of Florida
Gainesville, FL, USA
wangi@ufl.edu

Jesse Smith

Department of CISE
University of Florida
Gainesville, FL, USA
jd.smith@ufl.edu

Jaime Ruiz

Department of CISE
University of Florida
Gainesville, FL, USA
jaime.ruiz@ufl.edu

ABSTRACT

Prior work has shown that embodiment can benefit virtual agents, such as increasing rapport and conveying non-verbal information. However, it is unclear if users prefer an embodied to a speech-only agent for augmented reality (AR) headsets that are designed to assist users in completing real-world tasks. We conducted a study to examine users' perceptions and behaviors when interacting with virtual agents in AR. We asked 24 adults to wear the Microsoft HoloLens and find objects in a hidden object game while interacting with an agent that would offer assistance. We presented participants with four different agents: voice-only, non-human, full-size embodied, and a miniature embodied agent. Overall, users preferred the miniature embodied agent due to the novelty of his size and reduced uncanniness as opposed to the larger agent. From our results, we draw conclusions about how agent representation matters and derive guidelines on designing agents for AR headsets.

CCS CONCEPTS

• Human-centered computing → Human computer interaction (HCI) → Empirical studies in HCI

KEYWORDS

Augmented reality; embodied conversational agents

ACM Reference format:

Isaac Wang, Jesse Smith, and Jaime Ruiz. 2019. Exploring Virtual Agents for Augmented Reality. In *2019 CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, May 4–9, 2019, Glasgow, Scotland, UK. ACM, New York, NY, USA. Paper 281, 10 pages. <https://doi.org/10.1145/3290605.3300511>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CHI 2019, May 4–9, 2019, Glasgow, Scotland, UK

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-5970-2/19/05...\$15.00

<https://doi.org/10.1145/3290605.3300511>



Figure 1. Example of a user wearing an augmented reality headset and interacting with a virtual agent projected onto the real world through the headset.

1 INTRODUCTION

Augmented reality (AR) headsets are gaining popularity and beginning to enter the consumer market with the release of the Microsoft HoloLens [43], Meta 2 [42], and Magic Leap [36]. These headsets allow a user to see virtual objects or holograms projected onto a view of the real world. Unlike phone-based AR, which superimposes virtual images onto a camera feed, AR headsets are designed to operate entirely hands-free through an always-present, immersive view of virtual content, while keeping users situated in reality (Figure 1). This allows users to perform tasks in the real world without having to shift context in order to interact with a device. With these advantages, AR headsets have been touted as the technology that will eventually replace our smartphones [12,22].

One of the primary benefits of AR is the ability to provide a hands-free interface. Such interfaces are typically controlled through gesture and speech. For instance, the HoloLens AR headset uses gaze and tap gestures for interaction and also comes equipped with a version of Cortana, Microsoft's own speech-based assistant, specifically ported to AR [44]. These personal assistants or virtual agents, like those commonly found on smartphones (e.g., Siri [6], Google Assistant [26]), are a good fit for AR headsets with their ability to provide hands-free assistance. However, one common trait among current assistants is that, although they are becoming

more humanlike in terms of speech interaction capabilities, they lack a face and body.

Prior work has shown that an embodiment can be useful for a virtual agent [4,13,16,18,39], whether in the form of a humanoid avatar or some other kind of physical (or pseudo-physical) presence. Embodied Conversational Agents (ECAs) [17] have risen as a natural extension of dialogue systems, enriching the experience through a virtual avatar and adding a social aspect to the interaction [49,54]. Despite this, Cortana on the HoloLens does not have a body, instead remaining fully speech-driven. We see the potential for virtual agents in AR to take full advantage of embodiment, as the technology is already designed to display virtual objects in the real world.

We would expect the benefits of embodiment to carry over in the transition to AR, but would users prefer to work with an embodied or more traditional speech-only agent? We focus our research on understanding what visual representation (if any) users would prefer for agents in AR and how it affects their perceptions and behavior. In this paper, we present a study exploring users' interactions when engaging with four different agents in AR (voice-only, non-human, full-size embodied, and a miniature embodied agent) while completing a hidden object game. Our work provides the following contributions:

- (1) Quantitative results showing how an agent's visual representation affects users' gaze interaction behavior.
- (2) Qualitative results showing how visual representation affects users' perceptions of an agent.
- (3) Insights and implications for the design of virtual agents in augmented reality.

2 RELATED WORK

We focus our review of prior work on three categories: highlighting how ECAs are useful, emphasizing how having a physical form is preferred in human-robot interaction, and describing the current state of work on virtual agents for AR.

2.1 Embodied Conversational Agents

Embodied Conversational Agents [17], or ECAs, are dialogue systems that not only utilize a natural speech interface but also feature an embodiment, giving the agent a human (or occasionally non-human) appearance. The simple act of giving a system an embodiment can influence the way users view and use a system [52]. For instance, it allows a system to take on a more social role

[49,54], influencing users' perceptions and interactions as they treat the system more as another person [32,54].

With an embodiment, an agent can also leverage a full range of non-verbal communication to better engage with users. ECAs have been developed that leverage non-verbal communication to express emotion/affect [13,21] and to perform conversational functions [16,18,55]. For instance, Becker et al. [13] developed an ECA named Max, an agent with a full human body who would respond to and portray emotion. They found that including emotional capabilities helped increase perceptions of likeability and realism. Cassell and Thorisson [18] maintain that non-verbal behaviors for managing conversations [16] (e.g., signaling that the other person can speak) are more valuable than other behaviors used to convey emotion. Andrist et al. [4] showed how an agent's gaze can communicate information and help users with spatially oriented tasks. These studies show how there are advantages to embodiment other than simply providing a more likeable or engaging interface.

Certain behaviors and attributes of ECAs have also been shown to affect users' perceptions of those agents. For example, Kramer et al. [33] that natural behaviors such as eyebrow raising positively improved user evaluations. Mayer and DaPra [39] found that people performed better at learning and recalling information taught by an agent when it acted more humanlike in speech and gesture, and how social cues can impact learning outcomes. Similarly, Antos et al. [5] saw that users trusted agents that expressed emotion, but only if the emotions correctly matched what the agent was saying or doing. Overall, ECAs have the potential to change users' perceptions and behaviors.

Our work primarily explores ECAs in AR and looks at what forms they should take—from prior work, we know that ECAs affect how users respond due to the fact that they have an embodiment and can perform non-verbal functions. We may see this impact users' perceptions and interactions as they treat the agent more as another person.

2.2 Human-Robot Interaction

Research on embodiment from the field of human-robot interaction has also studied how having an embodiment or physical presence is beneficial. One key example is a study by Leyzberg et al. [34] in which users solved nonogram logic puzzles while interacting with a robot tutor. The tutor took on the form of either a disembodied voice, an onscreen agent, or a physically present robot. They found that having a physical form significantly

affected users' improvement in puzzle-solving times, which was attributed to the perceived authority afforded by a physical presence. This effect may also relate to perceived trust and social presence, as proposed by Bainbridge et al. [9]. If users perceive agents with a visual representation as more authoritative, trustworthy, or relatable, then an agent could influence users to better complete tasks.

Kiesler et al. [30] compared interaction with a physical robot against a virtual agent on a screen. The robot had an anthropomorphic design and could produce a range of facial expressions. Users found the robot more engaging due to its physicality. Similarly, Wainer et al. [59] studied having a physical robot as opposed to a telepresent robot and also found that the physical robot resulted in higher perceptions of the robot's watchfulness and that users rated their interaction with the robot as more enjoyable. These studies show promise for embodied agents in AR; as AR headsets are designed to display realistic holograms of virtual objects, we may see more of this physicality effect as opposed to just having an agent displayed on a screen.

2.3 Virtual Agents in Augmented Reality

Prior work has also looked into designing and developing agents for virtual and augmented reality systems. For instance, Barakonyi et al. [11] implemented an animation framework for handheld (e.g., smartphone-based) AR agents, making them fully animated and context-aware. Balcisoy et al. [10] detailed the creation of virtual humans in AR that both perceived and interacted with objects in the real world, for instance, to play checkers or place objects in the world. Likewise, Anabuki et al. [3] developed an agent that was designed to interact with both real and virtual worlds, actively responding to changes in both environments. These papers focused more on the design and implementation of agents in AR and less on understanding users' reactions to these agents.

Other researchers have focused on studying how users interact and respond to agents in AR. Miyake and Ito [47] investigated how users interacted with miniature agents that control specific objects, like a TV remote. In their study, users reported that it was easier to talk to agents that had a visual representation. Furthermore, Kim and Welch [31] laid out implications for agents' behaviors in AR. They described how the way an agent is situated in and interacts with the world can affect a user's experience in AR; for instance, an agent that is not well integrated with its environment may decrease a user's sense of presence.

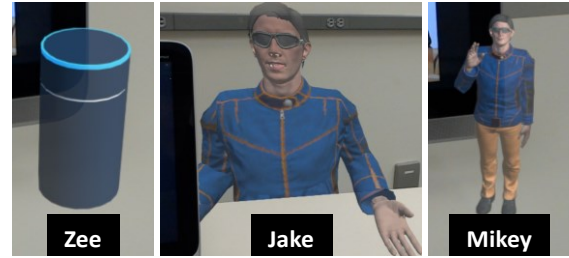


Figure 2. The three agents with visual representations, as viewed in context from the HoloLens. The fourth agent, Ava (not pictured), was a voice-only agent.

These exemplify how users can be influenced how agents act and behave in AR. We focus specifically on the visual representation an agent may assume, and investigate how an agent's representation changes users' perceptions and behaviors, and which agents do they prefer to work with in an AR setting.

3 AUGMENTED REALITY AGENTS

This section describes the different virtual agents that users would interact with while completing the hidden object game. We detail the agents' design, the hidden object game, supported user dialogue, and agent implementation.

3.1 Agents

We had four different agents for our users to interact with using the HoloLens while engaging in a hidden object game. Each agent had a different appearance (or lack thereof) and were designed to provide a range of visual representations for users to experience.

Ava (Voice-Only)

Ava was designed to primarily simulate the personal assistants of today (e.g., Siri). This agent responds to a range of speech commands (see following section), supporting the user as they complete the game on the computer. She does not have a physical appearance; rather, users would hear a disembodied female voice through the AR headset. Thus, our goal is that she would be able to assist users without distracting them away from the task.

Zee (Non-Human Representation)

This agent has a non-humanoid physical appearance. Zee was designed to mirror the appearance of a "smart home" device (e.g., Amazon Echo/Alexa), displaying as a cylindrical speaker that represents the agent in AR (Figure 2). The speaker sits to the lower-right next to a computer and is around 6 inches (15.2 cm) tall, the height of the Amazon Echo [2]. Zee behaves the same as Ava, the only

difference being the addition of a visual representation. Our goal is that this would provide a higher degree of realism and presence to the user, as Zee would have an actual appearance in the world as opposed to the voice-only Ava.

Jake (Full-Size Embodied Agent)

Jake is the first of our two humanlike embodied AR agents, designed to contrast with the prior two agents which mimic existing systems. This agent takes on the form of a full-sized human, visually represented as someone seated across from the user at the same desk (Figure 2). The use of an AR headset allows us to display Jake in full scale, calibrated to be at a size comparable to a physically present human (3.75 feet/1.14 meters tall from the floor, in a seated position).

As an embodied agent, Jake also exhibits a few non-verbal behaviors when interacting with the user. He displays non-verbal talking and waving gestures (for greetings, etc.) and also looks at the user when the user gazes at him, establishing mutual gaze [60,61]. Our goal is that these behaviors would allow Jake to be considered more realistic and relatable, possibly increasing users' perception of helpfulness and trust.

Mikey (Miniature Embodied Agent)

Our last agent, Mikey, is in all respects identical to Jake with the exception that he is a miniature version of Jake, the same height as the non-human agent Zee (Figure 2), and that he has a slightly higher-pitched voice than Jake to account for his smaller size. Mikey exhibits the same non-verbal behaviors as Jake, gesturing and gazing at the user when talking. Like Zee, Mikey's size makes him more suited to fit on a desk. Our goal is that Mikey would have the same benefits of embodiment as Jake, but in a smaller form that better fits the environment for desktop interactions, as he can stand on a desk next to a PC.

3.2 Hidden Object Game

We used hidden object puzzles as the task for users to complete while interacting with the agents. This was inspired from a study by Geven et al. [25] in which they asked people to find objects in a room while engaging with an agent for hints and other information. Our task simulates a visual search task and has the added benefit of not requiring the agent's help once the user knows what to look for. Thus, the agent would serve the role of an assistant, only providing help when asked. Additionally, such a task also explores the case where a user is visually focused and thus an agent with a visual representation could be a distraction.

For the individual hidden object tasks, we adapted puzzles from the I-Spy children's book series [37,38]. Each puzzle consisted of an image of a scene with multiple objects strewn about it. The goal for each puzzle was to find one object hidden in the scene, for example, the rabbit in Figure 3. Users interacted with the puzzles on a touchscreen PC, tapping the object once they found it in the image.

3.3 Supported Dialogue

Each agent accompanied the user as they worked to solve the hidden object puzzles. The agents would help the user find the object and provide hints when asked. For example, if the user asked, "What color is it," the agent would respond, "The rabbit is white." Agents supported the following speech intents for hints:

- **object:** States the object that the user must find.
- **color:** Describes the primary color(s) of the object.
- **shape:** Describes the object's specific shape or form.
- **context:** Describes objects that are near/around the goal object or other contextual clues.
- **location:** Describes the part of the image in which to look, such as a quadrant or half of the image.

Users did not know the possible hints or dialogue options but could interact naturally with the agents to find out. The agents would also list out hints if asked. Hints could either be requested by name ("I would like a color hint") or asked as a question ("what is near it?" or "is it blue?"), which added a degree of flexibility to the input prompts. Other intents were also included to support social dialogue such as "hello" and "thank you." All agents shared the same supported dialogue/hints, differing only in embodiment.

3.4 Implementation

We used the Unity game engine [57] and Microsoft's Mixed Reality Toolkit [45] to develop the agents for the HoloLens. Adobe Fuse [1] was used to generate the character models and animations for the two humanlike

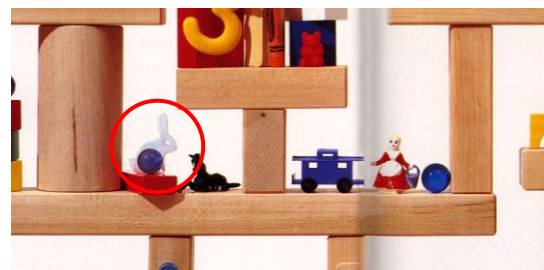


Figure 3. An excerpt from one of the hidden object puzzles. In this image, participants would need to find the rabbit (highlighted with a red circle).

agents. The Unity application functioned only as a client to display the agents and deliver the agents' speech to the user and did not contain any of the agent's intelligence, which was instead handled by the *dialogue manager* that ran on a separate PC as a server application.

The dialogue manager relied on Google's Dialogflow service [27], which provided automatic speech recognition and semantic intent extraction. User speech was sent to the service, which returned the speech intent (as described in the previous section). The application then generated an appropriate response or hint based on the currently active puzzle. The hidden object game was implemented in C# using the Windows Presentation Foundation (WPF) framework [46] and interfaced with the dialogue manager in order for the agent to know which puzzle the user was on and the hints available for the puzzle. The game ran on the same PC as the dialogue manager.

4 USER STUDY

We conducted a within-subjects study exploring how individuals interact with each of our four virtual agents while searching for objects in a hidden object game.

4.1 Participants

We recruited 24 participants from introductory computer science classes at a local university. Participants were adults between the ages of 18 and 28 (mean = 20.9, SD = 2.8), and 7 participants were female. Out of the 24 participants, 15 had prior experience with VR systems (e.g., Oculus Rift, HTC Vive), and only two specifically had experience with AR headsets (e.g., Microsoft HoloLens). All but one participant had experience with voice assistant technology due to the near ubiquity of assistants on smartphones. Participants received their choice of extra credit in a course or a \$10 Amazon gift card as compensation. Our study was approved by our Institutional Review Board.

4.2 Apparatus

Participants were seated in front of an HP Sprout [29] workstation (a 23-inch/58-cm touchscreen PC) and wore a Microsoft HoloLens headset (Figure 4). The hidden object game was displayed on the Sprout, and participants used the Sprout's touchscreen to interact with the game. The virtual agents were displayed on the HoloLens and positioned to the right of the workstation.

4.3 Procedure

At the start of the study, the researcher described the study and familiarized the participant with the HoloLens.

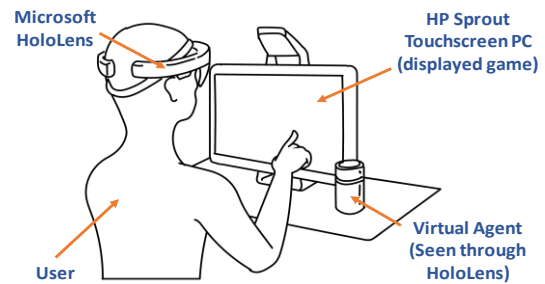


Figure 4. The experiment setup. Participants wore a Microsoft HoloLens and interacted with a hidden object game running on an HP Sprout PC.

The participant was then seated in front of the Sprout computer and put on the headset. The researcher then gave a brief summary of the hidden object game and introduced the participant to the first agent. The participant was then asked to complete a set of five hidden object puzzles, interacting naturally with the agent. The participant was asked to tap the object on the screen once they found it. After completing the set of puzzles, the participant filled out a short questionnaire describing their experience with the agent (see following section). This procedure was repeated for the other three agents; a different set of puzzles was used for each agent. Each participant interacted with all four agents, and order of exposure to the agents was counterbalanced using a balanced Latin square. At the end of the study, the participant was asked to describe which agent they preferred the most and which one they preferred the least before completing a few demographic questions.

We compiled four sets of puzzles, one set for each agent, with five puzzles in each set (20 total). Each set took on average eight minutes to solve as determined by a pilot test with four people (two male and two female; ages 21-23; mean = 21.5). The sets were initially randomized, and we kept the order of the puzzles the same for each participant, varying only the order agents appeared.

4.4 Virtual Agent Questionnaire

For the per-agent questionnaire, we asked participants to rate their experience with the agent along six visual analogue scales [23] rating different qualities of the agent: *helpfulness*, *presence*, *reliability*, *trust*, *distraction*, and *realism*. For example, "How would you rate the helpfulness of the agent?" These questions were adapted from prior work on embodied agents [8,14,28]. We also asked participants to rate how likely they would recommend the agent to a friend (*recommendation*) and how willing they would be to interact with the agent again (*continued use*) to get a general sense of user acceptance/satisfaction. We also asked participants to

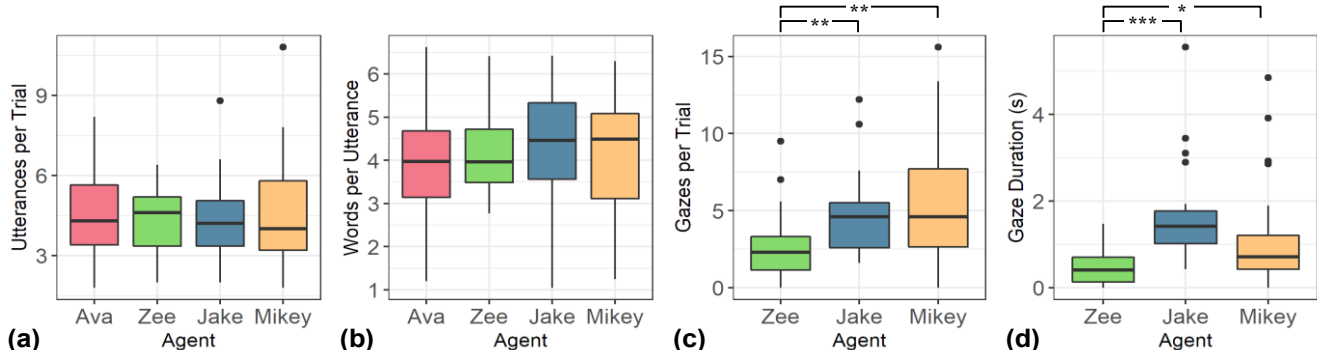


Figure 5. Boxplots comparing (a) number of utterances per trial, (b) words per utterance, (c) number of gaze events per trial, and (d) gaze duration across agents.

describe what they liked most and least about each agent, allowing us to capture detailed insights into their thoughts about the agents.

5 RESULTS

We report results from our study in two parts: the agent interaction data collected for each trial and the questionnaire responses collected for each agent.

5.1 Agent Interactions

For analysis, we logged data from each of the trials, including user utterances, utterance start/end times, agent utterances, times the user gazed at the agent, and incorrect taps on the puzzle (Figure 5). We omitted three trials (individual tasks, one from participant P8 with Mikey, P9 with Zee, and P20 with Jake) from analysis due to participants not being able to find an object and giving up, which forced the researcher to proceed to the next puzzle. This resulted in data from 24 participants * 20 trials per participant – 3 trials omitted = 477 total trials.

Task Completion Time

We analyzed the average time it took to complete a puzzle to see if participants completed tasks faster with a given

agent, since prior work [9,34,59] showed that users performed better with an embodied or physically-present agent/robot. The task completion time was measured from the time a user first saw a new puzzle to the time they found and tapped on the correct object. On average, participants took 67 seconds to complete a single puzzle. Table 1 lists the average task completion time for each agent. A within-subject repeated measures ANOVA (RM-ANOVA) revealed no significant effect of agent on puzzle completion time ($F_{3,60} = 2.11$, ns). Thus, overall, the visual representation of an agent did not have an impact on how fast participants completed a puzzle task.

Speech Interaction

We also looked at if participants spoke more to certain agents than to others. We wanted to see if people would engage or interact more with an agent based on their physical appearance. We recorded the total number of user utterances during each puzzle (Figure 5b) as well as the length of each utterance (Figure 5c). The average number of utterances and utterance lengths are reported in Table 1. RM-ANOVA revealed no significant effect of agent on the number of utterances ($F_{3,60} = 0.33$, ns) nor on utterance length ($F_{3,60} = 1.24$, ns), showing that participants did not speak significantly more or more often to any given agent.

Gaze Interaction

One of our goals was to see if having a human appearance would result in differences in users' non-verbal behavior, specifically users' gaze. For the three agents that had a visual representation, we looked at the number of times participants gazed at the agent (Figure 5d), as well as the duration of each gaze (Figure 5e). Gaze was estimated using head orientation based on the HoloLens headset, similar to in VR systems [7]. We determined gaze by seeing if the agent or agent's face was visible on screen. A single gaze event was counted as the user looking at the

	Ava (Voice)	Zee (Non-Human)	Jake (Full-Size)	Mikey (Mini)
Task Time (s)	73.8 (36.6)	61.1 (20.8)	70.0 (25.6)	64.1 (23.4)
Utterances per trial	4.5 (1.7)	4.3 (1.3)	4.4 (1.5)	4.5 (2.0)
Words per utterance	4.0 (1.3)	4.2 (1.1)	4.3 (1.3)	4.1 (1.3)

Table 1. Comparison of task time and utterance stats across the four agents. Standard deviations are in parentheses. No significant differences were observed.

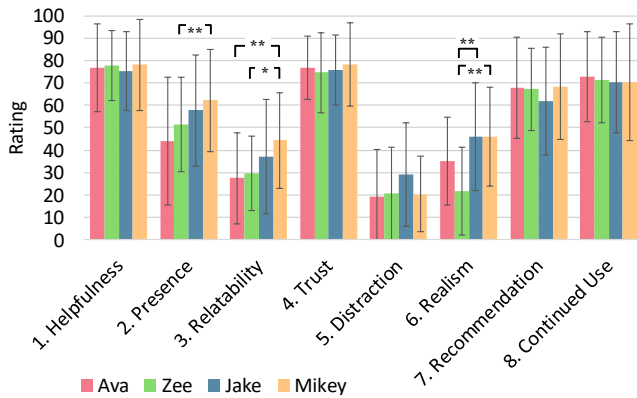


Figure 6. Graph of the average subjective ratings from the questionnaire. Each agent was rated against the eight scales. Error bars represent the standard deviation.

agent and then looking away. On average, Mikey had the largest number of gazes per trial (mean = 5.7 gazes, SD = 4.2), followed by Jake (mean = 4.8, SD = 2.6), and Zee (mean = 2.6, SD = 2.4). Jake had the longest average gaze duration (mean = 1.7 sec, SD = 1.1), followed by Mikey (mean = 1.2, SD = 1.3), and Zee (mean = 0.4, SD = 0.4).

RM-ANOVA revealed a significant main effect of agent on the number of gazes per trial ($F_{2,40} = 9.66$, $p < 0.001$) as well as on gaze duration ($F_{2,40} = 15.69$, $p < 0.001$). We did not observe any ordering or interaction effects. Post-hoc pairwise t-tests with Bonferroni correction showed that Zee (the non-human agent) had significantly fewer gazes than the other two humanlike agents Jake ($p < 0.01$) and Mikey ($p < 0.01$) and shorter gaze durations than both Jake ($p < 0.001$) and Mikey ($p < 0.05$).

We took a closer look at participants' gaze events to understand why users were gazing at an agent; was it in response to an agent speaking, was it when the user spoke to an agent, or was it simply to just look at the agent? Using the start/end times of each gaze event and utterance, we were able to determine the proportion of gazes that overlapped with utterances.

RM-ANOVA revealed an effect of agent on the proportion of gaze events that overlapped with both user and agent speech ($F_{2,40} = 18.28$, $p < 0.001$). For example, a user would gaze at an agent, speak, and then wait for its response before looking away. Post-hoc tests using Bonferroni correction revealed a significant difference between the three agents ($p < 0.05$ in all cases). Jake had the highest proportion of overlapped gaze events (0.25), Mikey with second-highest (0.13), and Zee with the lowest (0.05).

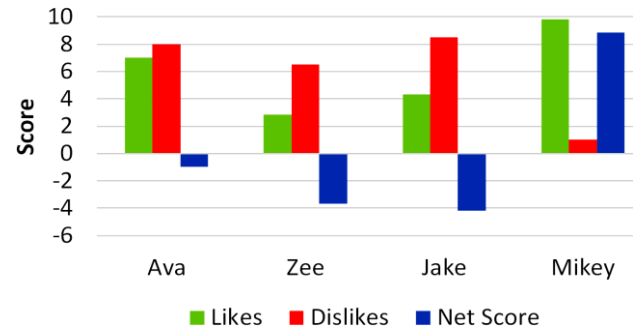


Figure 7. Participant ratings for most-liked and most-disliked agent. Net score (likes – dislikes) is also shown.

5.2 Questionnaire Responses

We analyzed participant's responses to understand their subjective perceptions and preferences for each agent.

Agent Perceptions

For analysis, we first quantified the responses on the eight visual analogue scales to a number from 0 to 100. We ran a repeated measures ANOVA on each of the eight scales individually in order to find differences between the four agents. Graphs comparing agent ratings for each of the scales are depicted in Figure 6.

We found a significant effect of agent representation on the presence ($F_{3,60} = 4.1$, $p < 0.01$), relatability ($F_{3,60} = 5.3$, $p < 0.01$), and realism ($F_{3,60} = 13.5$, $p < 0.001$) ratings. For the presence ratings, Mikey was rated as having significantly more presence than Zee ($p < 0.01$). For the relatability ratings, Mikey was rated as having significantly higher relatability than both Zee ($p < 0.05$) and Ava ($p < 0.01$). For the realism ratings, Jake and Mikey were both rated significantly higher than Zee ($p < 0.001$ for both). However, for all these ratings, we did not see a significant difference between Jake and Mikey. We did not find any significant effect of agent on any of the other scales (helpfulness, trust, distraction, recommendation, and continued use).

User Preference

To understand user preference for the four agents, we tallied up the participant votes for the agent they liked the most and liked the least. Some participants could not decide on a single agent, so their votes were divided into fractional votes. These results are highlighted in Figure 7.

Overall, Mikey (the miniature humanlike agent) had the greatest number of likes overall with 9.83 votes (out of 24 participants), with Ava (the voice-only agent) coming in second. Third was Jake, and last was Zee. In terms of most disliked agent, Jake was most disliked, followed

closely by Ava, and then Zee. Mikey had only one person who disliked him the most out of all four agents.

We also computed a *net score* for each agent by taking the number of likes and subtracting the number of dislikes (the third, blue bar in Figure 7). This allowed us to see which agents were liked universally and which had polarizing results based on individual user preference. Our results show that Mikey, the miniature human agent, was the most liked and the least disliked. On the other hand, Ava was almost equally liked and disliked, and Jake and Zee both had over twice as many dislikes than likes.

Agent Comments

To gain a fuller understanding of why certain agents were preferred over others, we looked at the comments from the open response questions. We analyzed the data using an affinity diagram to extract major themes. We note the main positive/negative themes for each individual agent below:

Ava (Voice-Only). Participants who liked Ava mentioned how she helped them focus on the task. For instance, one participant noted: “[Ava] was convenient because I never had anywhere to look away from and could strictly focus on the puzzle.” (P5) Ava’s voice-only nature also made it feel “like she was communicating from another room.” (P23) However, participants who disliked Ava the most said that they found her to be too similar to existing digital assistants: “it just seemed like a regular virtual agent similar to Siri.” (P7) Some disliked her voice-only nature: “the lack of a physical presence was odd” (P3).

Zee (Non-Human). Zee was liked for being more practical compared to the humanlike agents, as a simple representation without a full-on human appearance. Participants thought she gave helpful hints and liked how it was familiar to a real system (the Amazon Echo): “I think a lot more people would be comfortable with this agent simply because of how used to it we are.” (P14) The major downside to Zee was that she was considered the least humanlike and most robotic out of all the agents based on reported comments. “I felt like I was talking to a machine more than a helper/person.” (P2)

Jake (Full-Size). Participants who liked Jake emphasized his full-size figure as helping make him seem more humanlike. Additionally, participants liked his movements/behaviors and one specifically mentioned how those factors helped make him feel more “socially competent and more like a real person” (P2). Participants also liked that Jake had more presence than the other agents and how his appearance and location “made it feel more like I was having a conversation with someone rather

than looking down or demanding them” (P22). However, others strongly disliked Jake’s appearance, primarily disliking how large he was. His size made him feel “off-putting” (P14) and “uncanny” (P13). One participant (P10) best described this as “having a huge animated guy right next to you was a bit much.”

Mikey (Miniature). Mikey was liked most for feeling natural when interacting with him and he was also considered very humanlike. Participants noted how his small size made him feel more approachable and relatable: “it was almost cute and like a friend due to size but relatability [sic]” (P13). Mikey was also commonly liked due to the novelty factor of having a small person as a personal assistant, with some calling him entertaining and amusing: “he felt like a little helper that was in the room with me, wasn’t too distracting, and was amusing because he was so small.” (P15) A few noted how Mikey could still be more lifelike: “While he was more lifelike than other virtual agents I’ve seen, he was still relatively wooden/seemingly emotionless.” (P15)

6 DISCUSSION

In this section, we discuss the results of our study, relating it to prior work and specifically focusing on how our results can be best applied to designing virtual agents for AR systems. We start by discussing user preference for the miniature embodied agent Mikey, before analyzing how our users perceived the other agents and what implications it has for agents in AR.

6.1 Embodied Agents for AR

We found that the two humanlike embodied agents, Mikey and Jake, resulted in higher numbers of gazes than the non-human agent Zee. Participants also rated the miniature Mikey as the most-liked agent overall, primarily because they found him more novel, but also rated him as having higher presence and relatability when compared with the voice-only Ava. Results were mainly the same for Jake in terms of interaction and comments; the main difference came down to participants reporting that the small form made Mikey more relatable and humanlike.

These results align with prior work on embodied agents outside of AR. A study by Wagner et al. [58] found that their embodied agents were rated as more likeable and engaging than those without embodiments. They also reported a novelty effect: their embodied agents were rated as more enjoyable and fun than a disembodied voice. Similarly, Wainer et al. [59] also saw a novelty effect, but across all their embodiments. In contrast, we saw that Mikey had participants mention how he was amusing and

novel while the other agents did not invoke such responses.

Examination of gaze interaction showed that participants gazed at the two humanlike embodied agents more often than the non-human Zee, looking at the agents both when speaking and being spoken to by the agent. We believe this to be because people naturally gaze/fixate on faces [15,19]. Additionally, the mutual gaze that the humanlike agents provided may have encouraged users to gaze at them more. Takeuchi and Naito [54] saw similar gaze behavior, at least anecdotally, where users would gaze more at a humanlike agent than an agent with a more abstract representation. The fact that participants gazed more at these agents shows promise for improving agent interaction through non-verbal communication, which prior work has shown to benefit interaction [4,13,18]. Our embodied agents have the ability to deliver whatever non-verbal communication they need (through gestures, expressions, gaze, etc.)—if users did not look at the agent, then it would be difficult for an agent to convey any sort of non-verbal information.

Overall, we did not produce any strong evidence against having embodiments for agents in AR, which shows promise for being able to adapt current work on ECAs to AR systems while retaining the same benefits. This is not to say that there is no benefit to having a voice-only agent; our Ava agent was second-most liked, with participants touting her ability to keep them focused on the task at hand and how she was less distracting. Based on the task, an agent's role may change and subsequently a user's interaction with the agent may also differ. More research could be conducted with tasks that support such as manipulating 3D objects in AR or helping with real-world applications (such as an assembly line). However, based on our results, we recommend the use of a miniature embodied agent in AR, at least for desktop-oriented tasks.

6.2 Size and Uncanniness

From our results, we saw that participants preferred the smaller Mikey over the full-size Jake agent. Based on the quantitative results of the interaction and the questionnaire, we did not see any significant differences between Jake and Mikey, apart from gaze interaction. However, when participants had to choose between agents, Mikey was more liked than Jake. Although participants found Jake realistic and humanlike, he was mentioned as being too large and imposing, with comments about his appearance and how he felt weird or creepy as he sat there. Some liked Jake because he was

very humanlike, while others disliked him because he was *too* humanlike.

We believe this to be a result of Jake falling into the “uncanny valley,” first proposed by Mori et al. [48], where the likeability of a robot (or virtual agent in our case) dips down as the robot looks and acts more like a human but not quite enough. This is a particular issue for computer generated avatars and is a difficult obstacle to identify and overcome [56]. The uncanny valley can be affected by different attributes [53]. For instance, the original theory by Mori et al. [48] mentioned both appearance and movement as factors, and MacDorman et al. [35] noted how even the proportions of facial features could impart a sense of eeriness. Thus, the uncanniness of Jake may be a result of his specific avatar characteristics, whether it be the naturalness of his movements/animation or specific design features related to his appearance.

However, in our study, Jake had the same appearance and movements as Mikey apart from size, but participants still mentioned that Mikey felt more natural/humanlike as well as relatable and approachable. A smaller agent would be considered less human and help keep it from falling into the uncanny valley. Thus, the uncanny valley may also be strongly affected by size, a factor that has not been sufficiently explored in prior work. Thus, a smaller agent may be preferred as it still provides the expressiveness of an ECA while also avoiding the uncanniness that a full-size human agent may incur.

6.3 Representation and Anthropomorphism

We saw that our participants felt that the non-human agent Zee was the least realistic and humanlike based on their comments in the questionnaire, although we did not see a significant difference in the quantitative results. Zee had the most comments out of all the agents regarding how she felt more robotic and less personable. This was primarily due to her visual representation: “*Since it wasn’t a person but an inanimate object it felt less like I was interacting with somebody.*” (P23) On the other hand, P8 felt that this was acceptable inasmuch as the agent was not *trying* to be human: “*I think this is sort of fine because it doesn’t feel weird, [Zee] is not trying to be a human to me whereas Jake was.*” Both agents were acceptable but ended up filling different social roles based on their appearance (one as a human and one as a robot).

We expected Zee to feel less human than the explicitly humanlike agents, but we did not expect to see more comments about how robotic Zee was compared with the voice-only Ava. The only difference between Ava and Zee was in their appearance. However, having the appearance

of a smart device strongly solidified the concept of Zee as not being human, more so than having no appearance.

We believe this to be related to the tendency for people to treat computers as if they were people, from a social interaction perspective [52]. For Ava, users did not see any visuals representing the voice-only agent, so they were able to still consider the voice as just a person talking. This supports the theory presented by Nowak and Biocca [51] where users would likely assume a “default image” for the agent, considering it human until they were presented with information to the contrary. In our study, participant P2 best described this effect:

“I did not see its face as it helped me. I instead saw a cylinder like the devices that exist in real life. The first assistant had only a voice, which allowed me to imagine that it had a face and thoughts that were just not shown.”

Thus, we caution designers to avoid giving an agent a non-human visual representation similar to what we used. Unless the goal is to make users feel like they are interacting with a robot, this has the potential to make an agent feel more robotic and less personable.

7 LIMITATIONS AND FUTURE WORK

In our study, we only looked at four different agent representations. There are other appearances that an agent may assume, such as a cartoon-like agent or an anthropomorphic robot. Other factors such as voice, movement, and location could also have an effect on users. We also did not look at if the gender of the virtual agent had an impact on users’ behaviors and perceptions. Prior work has shown some preference to female embodied agents [24,41,50,62], while others have found no difference in preference [20,32] or a preference for male agents [40]. However, these studies were specifically looking at factors such as credibility or likeability, which was not our research emphasis. Our results show that embodied humanlike agents are effective for AR, with more gaze interactions and perceptions of presence and relatability. Additionally, we saw that Ava (a female voice agent) was rated second-highest in terms of preference, even more than Jake. However, future work would need to probe the design space and tease apart how these factors would further influence users’ preferences for agents in AR.

Another limitation of our work is the inherently small field of view (FoV) of the HoloLens display, which does not encompass a user’s entire peripheral range. This could have had an impact on their sense of presence or the realism of an agent. In our study, participants could see Mikey and Zee fully, but only the head and shoulders of

Jake. This may also have had an effect on perceptions of Jake, but participants were more concerned about his uncanny size than being able to see him. Hardware with a higher FoV could be used to see if FoV affects a users’ sense of presence.

We would also look into replicating our study with different scenarios to understand more about how the choice of task affects users’ preferences for an agent. In terms of the tasks we used for the study, we focused on simulating a scenario based from prior work by Geven et al. [25] where the user is occupied with a visual task and the agent acts as an assistant. Users may have different perceptions and preferences for agents based on the task they have to complete; for instance, driving a car may be too critical to afford any possible distractions. Additionally, we could also explore how a smaller agent would fare in tasks where an agent’s advice would be crucial in completing tasks (such as in teaching scenarios) and understand if the smaller size affects an agent’s perceived capabilities or authority.

8 CONCLUSION

In this paper, we presented results from a study exploring user preferences, behaviors, and perceptions when interacting with different virtual agents in augmented reality. We showed how users gazed more at the humanlike agents than a non-human agent, and how users would gaze at an agent while speaking to it and while it speaks to them. Additionally, we showed that, overall, users preferred the miniature embodied agent Mikey, as he was rated higher with regards to presence, relatability, and realism. We also described how Mikey’s size played into a reduced feeling of uncanniness compared against the full-size embodied agent Jake, and how the presence of a non-human embodiment affected how users felt about it. Our results help direct the design of virtual agents for AR headsets and motivate future research on developing virtual agents for different AR applications.

ACKNOWLEDGMENTS

This work is partially supported by the U.S. Defense Advanced Research Projects Agency (DARPA) and the U.S. Army Research Office (ARO) under contract #W911NF-15-1-0459, and the National Science Foundation Grant Award #IIS-1750840. Opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect these agencies’ views.

REFERENCES

- [1] Adobe. 2018. Adobe Fuse. Retrieved September 5, 2018 from <https://www.adobe.com/products/fuse.html>
- [2] Amazon. 2018. Amazon Echo. Retrieved September 20, 2018 from <https://www.amazon.com/all-new-amazon-echo-speaker-with-wifi-alexa-dark-charcoal/dp/B06XCM9LJ4>
- [3] Mahoro Anabuki, Hiroyuki Kakuta, Hiroyuki Yamamoto, and Hideyuki Tamura. 2000. Welbo: An Embodied Conversational Agent Living in Mixed Reality Space. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '00), 10–11. <https://doi.org/10.1145/633292.633299>
- [4] Sean Andrist, Michael Gleicher, and Bilge Mutlu. 2017. Looking Coordinated: Bidirectional Gaze Mechanisms for Collaborative Interaction with Virtual Characters. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17), 2571–2582. <https://doi.org/10.1145/3025453.3026033>
- [5] Dimitrios Antos, Celso De Melo, Jonathan Gratch, and Barbara J. Grosz. 2011. *The Influence of Emotion Expression on Perceptions of Trustworthiness in Negotiation*. AAAI Press. Retrieved August 28, 2018 from <https://dash.harvard.edu/handle/1/5344528>
- [6] Apple. 2018. iOS - Siri. Retrieved September 4, 2018 from <https://www.apple.com/ios/siri/>
- [7] Rowel Atienza, Ryan Blonna, Maria Isabel Saldares, Joel Casimiro, and Vivencio Fuentes. 2016. Interaction techniques using head gaze for virtual reality. In *2016 IEEE Region 10 Symposium (TENSYP)*, 110–114. <https://doi.org/10.1109/TENCONSpring.2016.7519387>
- [8] Jeremy N Bailenson, Eyal Aharoni, Andrew C Beall, Rosanna E Guadagno, Aleksandar Dimov, and Jim Blascovich. 2004. Comparing behavioral and self-report measures of embodied agents' social presence in immersive virtual environments. In *Proceedings of the 7th Annual International Workshop on PRESENCE*, 1864–1105.
- [9] Wilma A. Bainbridge, Justin W. Hart, Elizabeth S. Kim, and Brian Scassellati. 2011. The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. *International Journal of Social Robotics* 3, 1: 41–52. <https://doi.org/10.1007/s12369-010-0082-7>
- [10] Selim Balcisoy, Marcelo Kallman, Rémy Torre, Pascal Fua, and Daniel Thalmann. 2002. Interaction techniques with virtual humans in mixed environments. In *5th IEEE EMBS International Summer School on Biomedical Imaging*, 2002., 6 pp.-. <https://doi.org/10.1109/SSBI.2002.1233993>
- [11] Istvan Barakonyi and Dieter Schmalstieg. 2006. Ubiquitous Animated Agents for Augmented Reality. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality* (ISMAR '06), 145–154. <https://doi.org/10.1109/ISMAR.2006.297806>
- [12] Dina Bass and Mark Gurman. 2017. Microsoft Gets Serious About Hardware. *Bloomberg.com*. Retrieved September 4, 2018 from <https://www.bloomberg.com/news/features/2017-05-03/microsoft-gets-hardware-religion>
- [13] Christian Becker, Stefan Kopp, and Ipke Wachsmuth. 2007. Why emotions should be integrated into conversational agents. *Conversational informatics: an engineering approach*: 49–68.
- [14] Elisabetta Bevacqua, Romain Richard, and Pierre De Loo. 2017. Believability and Co-presence in Human-Virtual Character Interaction. *IEEE Computer Graphics and Applications* 37, 4: 17–29. <https://doi.org/10.1109/MCG.2017.3271470>
- [15] Markus Bindemann, A. Mike Burton, Ignace T. C. Hooge, Rob Jenkins, and Edward H. F. de Haan. 2005. Faces retain attention. *Psychonomic Bulletin & Review* 12, 6: 1048–1053. <https://doi.org/10.3758/BF03206442>
- [16] Justine Cassell, Timothy Bickmore, Mark Billinghurst, Lee Campbell, Ken Chang, Hannes Vilhjálmsón, and Hao Yan. 1999. Embodiment in Conversational Interfaces: Rea. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '99), 520–527. <https://doi.org/10.1145/302979.303150>
- [17] Justine Cassell, Joseph Sullivan, Elizabeth Churchill, and Scott Prevost. 2000. *Embodied conversational agents*. MIT press, Cambridge, Mass.
- [18] Justine Cassell and Kristinn R. Thorisson. 1999. The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence* 13, 4–5: 519–538. <https://doi.org/10.1080/088395199117360>
- [19] Moran Cerf, E. Paxon Frady, and Christof Koch. 2009. Faces and text attract gaze independent of the task: Experimental data and computer model. *Journal of Vision* 9, 12: 10.1–15. <https://doi.org/10.1167/9.12.10>
- [20] Andrew J. Cowell and Kay M. Stanney. 2003. Embodiment and Interaction Guidelines for Designing Credible, Trustworthy Embodied Conversational Agents. In *Intelligent Virtual Agents* (Lecture Notes in Computer Science), 301–309.
- [21] Marc Fabri, Salima Y. Awad Elzouki, and David Moore. 2007. Emotionally Expressive Avatars for Chatting, Learning and Therapeutic Intervention. In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments* (Lecture Notes in Computer Science), 275–285. https://doi.org/10.1007/978-3-540-73110-8_29
- [22] Facebook. 2017. F8 2017: AI, Building 8 and More Technology Updates From Day Two. Retrieved September 4, 2018 from <https://newsroom.fb.com/news/2017/04/f8-2017-day-2/>
- [23] Darren Flynn, Paul van Schaik, and Anna van Wersch. 2004. A Comparison of Multi-Item Likert and Visual Analogue Scales for the Assessment of Transactionally Defined Coping Function1. *European Journal of Psychological Assessment* 20, 1: 49–58. <https://doi.org/10.1027/1015-5759.20.1.49>
- [24] Jodi Forlizzi, John Zimmerman, Vince Mancuso, and Sonya Kwak. 2007. How Interface Agents Affect Interaction Between Humans and Computers. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (DPPI '07), 209–221. <https://doi.org/10.1145/1314161.1314180>
- [25] Arjan Geven, Johann Schrammel, and Manfred Tscheligi. 2006. Interacting with Embodied Agents That Can See: How Vision-enabled Agents Can Assist in Spatial Tasks. In *Proceedings of the 4th Nordic Conference on Human-computer Interaction: Changing Roles* (NordiCHI '06), 135–144. <https://doi.org/10.1145/1182475.1182490>
- [26] Google. 2018. Google Assistant. Retrieved September 4, 2018 from <https://assistant.google.com/>
- [27] Google. 2018. Dialogflow. Retrieved September 5, 2018 from <https://dialogflow.com/>
- [28] Jonathan Gratch, Ning Wang, Jillian Gerten, Edward Fast, and Robin Duffy. 2007. Creating Rapport with Virtual Agents. In *Intelligent Virtual Agents* (Lecture Notes in Computer Science), 125–138.
- [29] HP. 2018. HP Sprout. Retrieved September 5, 2018 from <https://store.hp.com/us/en/mdp/sprout-by-hp-349520--1>
- [30] Sara Kiesler, Aaron Powers, Susan R. Fussell, and Cristen Torrey. 2008. Anthropomorphic Interactions with a Robot and Robot-like Agent. *Social Cognition* 26, 2: 169–181. <https://doi.org/10.1521/soco.2008.26.2.169>
- [31] Kangsoo Kim and Greg Welch. 2015. Maintaining and Enhancing Human-Surrogate Presence in Augmented Reality. In *2015 IEEE International Symposium on Mixed and Augmented Reality Workshops*, 15–19. <https://doi.org/10.1109/ISMARW.2015.13>
- [32] T. Koda and P. Maes. 1996. Agents with faces: the effect of personification. In *Proceedings 5th IEEE International Workshop on Robot and Human Communication. RO-MAN'96 TSUKUBA*, 189–194. <https://doi.org/10.1109/ROMAN.1996.568812>
- [33] Nicole C. Krämer, Nina Simons, and Stefan Kopp. 2007. The Effects of an Embodied Conversational Agent's Nonverbal Behavior on User's Evaluation and Behavioral Mimicry. In *Proceedings of the 7th International Conference on Intelligent Virtual Agents* (IVA '07), 238–251. https://doi.org/10.1007/978-3-540-74997-4_22
- [34] Daniel Leyzberg, Samuel Spaulding, Mariya Toneva, and Brian Scassellati. 2012. The physical presence of a robot tutor increases cognitive learning gains. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.
- [35] Karl F. MacDorman, Robert D. Green, Chin-Chang Ho, and Clinton T. Koch. 2009. Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior* 25, 3: 695–710. <https://doi.org/10.1016/j.chb.2008.12.026>

- [36] Magic Leap, Inc. 2018. Magic Leap. Retrieved September 4, 2018 from <https://www.magicleap.com/>
- [37] Jean Marzollo. 1992. *I spy: a book of picture riddles*. Scholastic Inc, New York.
- [38] Jean Marzollo. 1993. *I spy mystery: a book of picture riddles*. Scholastic, New York.
- [39] Richard E. Mayer and C. Scott DaPra. 2012. An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied* 18, 3: 239–252. <https://doi.org/10.1037/a0028616>
- [40] Helen McBreen and Mervyn Jack. 2001. Evaluating humanoid synthetic agents in e-retail applications. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 31, 5: 394–405. <https://doi.org/10.1109/3468.952714>
- [41] Helen McBreen and Mervyn Jack. 2000. Empirical evaluation of animated agents in a multi-modal e-retail application. In *Proc. AAAI Fall Symposium: Socially Intelligent Agents*, 122–126.
- [42] Meta Company. 2018. Meta Augmented Reality. Retrieved September 4, 2018 from <https://www.metavision.com/>
- [43] Microsoft. 2018. HoloLens. Retrieved September 4, 2018 from <https://www.microsoft.com/en-us/hololens>
- [44] Microsoft. 2018. Cortana on HoloLens. Retrieved September 4, 2018 from <https://support.microsoft.com/en-us/help/12630/hololens-cortana-on-hololens>
- [45] Microsoft. 2018. MixedRealityToolkit - Unity. Retrieved September 5, 2018 from <https://github.com/Microsoft/MixedRealityToolkit-Unity>
- [46] Microsoft. 2018. What is WPF? Retrieved September 20, 2018 from <https://docs.microsoft.com/en-us/visualstudio/designers/getting-started-with-wpf>
- [47] Shinji Miyake and Akinori Ito. 2012. A spoken dialogue system using virtual conversational agent with augmented reality. In *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, 1–4.
- [48] Masahiro Mori. 2012. The Uncanny Valley. *IEEE Spectrum: Technology, Engineering, and Science News*. Retrieved October 20, 2016 from <http://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>
- [49] Clifford Nass, Jonathan Steuer, and Ellen R. Tauber. 1994. Computers Are Social Actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '94), 72–78. <https://doi.org/10.1145/191666.191703>
- [50] Andreea Niculescu, Dennis Hofs, Betsy van Dijk, and Anton Nijholt. 2010. *How the Agent's Gender Influence Users' Evaluation of a QA System*. Social Science Research Network, Rochester, NY. Retrieved September 18, 2018 from <https://papers.ssrn.com/abstract=2954317>
- [51] Kristine L. Nowak and Frank Biocca. 2003. The Effect of the Agency and Anthropomorphism of Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoper. Virtual Environ.* 12, 5: 481–494. <https://doi.org/10.1162/10547460332276128>
- [52] Byron Reeves and Clifford Ivar Nass. 1996. *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge university press.
- [53] Michael Seymour, Kai Riemer, and Judy Kay. 2017. Interactive Realistic Digital Avatars - Revisiting the Uncanny Valley. *Hawaii International Conference on System Sciences 2017 (HICSS-50)*. Retrieved from <https://aisel.aisnet.org/hicss-50/cl/hci/4>
- [54] Akikazu Takeuchi and Taketo Naito. 1995. Situated Facial Displays: Towards Social Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '95), 450–455. <https://doi.org/10.1145/223904.223965>
- [55] Kristinn R Thórisson. 1997. Gandalf: An embodied humanoid capable of real-time multimodal dialogue with people. In *Agents*, 536–537.
- [56] Angela Tinwell. 2009. Uncanny As Usability Obstacle. In *Proceedings of the 3D International Conference on Online Communities and Social Computing: Held As Part of HCI International 2009* (OCSC '09), 622–631. https://doi.org/10.1007/978-3-642-02774-1_67
- [57] Unity Technologies. 2018. Unity. Retrieved September 7, 2018 from <https://unity3d.com>
- [58] Daniel Wagner, Mark Billinghurst, and Dieter Schmalstieg. 2006. How Real Should Virtual Characters Be? In *Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology* (ACE '06). <https://doi.org/10.1145/1178823.1178891>
- [59] Joshua Wainer, David J. Feil-Seifer, Dylan A. Shell, and Maja J. Mataric. 2007. Embodiment and Human-Robot Interaction: A Task-Based Perspective. In *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, 872–877. <https://doi.org/10.1109/ROMAN.2007.4415207>
- [60] Ning Wang and Jonathan Gratch. 2010. Don'T Just Stare at Me! In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10), 1241–1250. <https://doi.org/10.1145/1753326.1753513>
- [61] Yuichiro Yoshikawa, Kazuhiko Shinozawa, Hiroshi Ishiguro, Norihiro Hagita, and Takanori Miyamoto. 2006. Responsive robot gaze to interaction partner. Retrieved September 16, 2018 from <http://www.roboticsproceedings.org/rss02/p37.html>
- [62] John Zimmerman, Ellen Ayoub, Jodi Forlizzi, and Mick McQuaid. 2005. Putting a face on embodied interface agents. *Proc. of Designing Pleasurable Products and Interfaces*. Eindhoven Technical University Press: 233–248.