Exploring Users' Perceptions on Position, Gaze Direction, and Gender of Virtual Agents in Augmented Reality

Rodrigo L. Calvo
rodrigo.calvo@ufl.edu
University of Florida
Gainesville, Florida, USA

Heting Wang
heting.wang@ufl.edu
University of Florida
Gainesville, Florida, USA

Alexander Barquero
a.barquero@ufl.edu
University of Florida
Gainesville, Florida, USA

Jaime Ruiz
jaime.ruiz@ufl.edu
University of Florida
Gainesville, Florida, USA

ABSTRACT

Prior research has highlighted users' preferences for embodiment when interacting with virtual agents in augmented reality headsets. However, open questions remain regarding users' preferences towards agent placement and gaze direction. In our study, we asked 48 adults to wear the Microsoft HoloLens 2 and find objects in a hidden object game with the help of embodied agents. We examined four distinct agent configurations for both male and female agents: a human-size agent standing beside participants, a human-size agent sitting beside participants, a small desk agent facing the screen, and a small desk agent facing the participant. Overall, participants preferred male over female virtual agents when receiving assistance, and no consistent preference emerged regarding the agents' position or gaze direction. From our results, we build upon existing guidelines for designing better virtual agents for AR with headsets.

Figure 1: Virtual agents used in the study. Top row: male agents; bottom row: female agents. From left to right: Nova (small-sized, table, facing the participant), Drew (small-sized, table, facing the screen), Kai (human-sized, sitting beside the participant), and Angel (human-sized, standing beside the participant).

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

augmented reality, embodied virtual agents, agent placement, agent gender

ACM Reference Format:
Rodrigo L. Calvo, Heting Wang, Alexander Barquero, and Jaime Ruiz. 2018. Exploring Users' Perceptions on Position, Gaze Direction, and Gender of Virtual Agents in Augmented Reality. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 10 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Augmented reality (AR) headsets have become increasingly popular in recent years. They have caught the attention and interest of a wide audience since they offer a unique and immersive way to experience digital content in the real world. AR devices, such as the Microsoft Hololens 2 [40] and Magic Leap 2 [24], allow users to place or superimpose virtual objects, artifacts, or holograms over what is seen in the real world. This enhances how users see and interact with their surroundings. These AR headsets are becoming popular not just because they are appealing, but also because they have the potential to revolutionize sectors like health [1] and education [14].

AR devices offer versatile user interactions, going beyond keyboards and mice with natural gestures and speech. Additionally, as AR technology continues to advance, the concept of virtual agents has emerged as another interactive modality. Virtual agents interpret users' speech, movements, or both using sensors, but may lack

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

a physical form or representation. Research has shown that giving a virtual agent a physical presence, whether as a human avatar or another form of embodiment, can be beneficial [2, 7, 9, 33, 52]. For example, embodied conversational agents (ECAs) enhance a dialogue system experiences by adding a social aspect to the interaction with the virtual agent [44, 48]. Further research, such as Wang et al.'s [54] study, has explored users' preferences towards different embodied agents' features in AR environments, revealing a preference for human-like agents.

Building on the findings about visual representation preferences, the design attributes of virtual agents, such as gender, position, and gaze direction have also attracted attention in research. While studies like Wang et al. [54] have shown preferences for human-like agents in augmented reality settings, they did not explore gender dynamics. Existing literature presents mixed findings regarding gender preferences for virtual agents, with some studies favoring female representations [15, 34, 45, 58], others favoring male representations [35], and some indicating no significant difference [12, 21]. This diversity of findings highlights the need to investigate the role of gender alongside other design attributes in shaping user perceptions and preferences towards embodied agents. In terms of positioning and gaze direction, a significant portion of research has been dedicated to positioning virtual agents in AR in ways that ensure they do not obstruct objects and are in harmony with the environment [23, 56]. However, these studies often do not emphasize a human-centric approach, such as positioning agents side-by-side with users, as one would in real-life human interactions [19]. Our study investigates whether an AR virtual agent's position relative to the user, gaze direction, and gender affects the users' behaviors and perceptions of the agent. Our hypothesis is that the side-by-side agent configuration, where the agent stands or sits beside the user, mirrors a more natural human interaction, similar to how people typically assist one another [5]. In contrast, conventional EVAs have been designed to face the user when interacting with them [15, 22, 54]. To this end, we aim to answer the following question: What is the impact of an AR virtual agent's position, gaze direction, and gender on the users' behaviors and perceptions during an AR activity? In this paper, we present the results and findings from studying users' interactions with AR virtual agents while completing four sets of find-the-hidden-object puzzles (see section 3.3). Each participant engaged with four uniquely positioned agents, as shown in Figure 1. Participants were randomly split into two groups: half interacted with male versions of these agents, while the other half interacted with their female counterparts.

Our results revealed that while agent positioning did not significantly alter user perceptions, gender biases were evident. Male agents outperformed their female counterparts in areas such as appropriateness, willingness, and likability. From our qualitative analysis, Nova (small desk agent facing the participant) emerged as a preferred agent due to its positioning, while Angel (human-size agent standing beside the participant), despite certain positive feedback, was the least favored, often described as intrusive. Our work provides the following contributions:

(1) Quantitative results highlighting gender biases in user perceptions of virtual agents in augmented reality.
(2) Qualitative insights into participants' subjective preferences, revealing specific likes and dislikes associated with each agent's design, voice, and positioning.
(3) Implications and recommendations for the design of virtual agents in augmented reality taking position and gender into consideration.

2 RELATED WORK

In this section, we present the domain of embodied virtual agents, highlighting both the advancements and open questions in research surrounding virtual agent positioning and gender.

2.1 Embodied Virtual Agents

Embodiment has emerged as a highly effective approach for presenting interactive virtual agents to users, contributing to enhanced user-agent interactions and immersive experiences [20, 28, 29, 36]. Researchers have established contextual design guidelines for creating embodied virtual agents (EVAs) [3, 46], focusing on agents that can engage users in a more human-like manner.

Wang et al. [54] conducted research in the realm of EVAs' appearance within headset-enabled AR environments. By studying various methods to generate the visual attributes of these agents, they aimed to determine the ideal design choices based on users' preferences, behaviors, and perceptions. The findings highlighted the significance of agent design in fostering rapport and trust between users and virtual entities, revealing the importance of creating visually appealing and relatable virtual agents.

Trust is also a crucial factor that underpins user-agent interactions, particularly in the context of EVAs. Morinezhad and Solovey [42] explored the intricate dynamics of trust in agents, emphasizing its relative nature. The level of trust users placed in EVAs is contingent upon both their past interaction experiences with such agents and the agents' current behavior. Their research outlined that integrating cooperative behavior into EVAs fosters higher levels of perceived trustworthiness, enhancing the overall user experience.

In the field of education, EVAs have shown promising outcomes. Grivokostopolou et al. investigated the impact of embodied pedagogical agents in virtual learning environments [17]. Their study demonstrated that the presence of embodied agents significantly improved students' learning experiences, leading to heightened engagement and more effective knowledge-creation processes. These findings demonstrate the potential of EVAs as valuable educational tools.

In summary, embodiment stands as a compelling and effective paradigm for virtual agent interaction. Studies in this area have elucidated key design considerations for crafting EVAs that align with users' preferences and contextual requirements. The appearance of embodied agents can play a pivotal role in fostering rapport and trust with users. Additionally, understanding the dynamic nature of users' trust in EVAs, as well as the positive impact of cooperative behavior, can contribute to more successful user-agent interactions.

2.2 Virtual Agent Positioning

Virtual agent positioning refers to the placement and orientation of an agent with respect to the user and entities within the virtual environment. When designing virtual agents for AR environments, it is crucial to consider the positioning of the agent to simulate natural human interactions. For example, Kendon [19] discussed the spacing and orientation in human co-present interactions, shedding light on how people naturally adjust their positioning based on their interpersonal connections. He introduced various positioning arrangements (e.g., circular, side-by-side) and discussed how the group maintained spatial and orientational relationships while leaving room for group activities. These findings provided valuable insights into guiding the design of human-agent interactions.

Various factors can impact user preferences for agent positioning, including the agent's position regarding the objects [23, 56], the user [47, 56], and the user's objective in the task [49, 56]. These studies computed the position of a virtual agent as the outcome of a model incorporating the user and environmental factors. Lang et al. [23] employed scene semantics to guide the positioning of a virtual agent in mixed reality. The position of their agent is computed using a cost function of the positions of the key objects in a 3D environment. Users favored the agent's positioning (e.g., location, orientation, and overall impression) than placing it randomly within a 30-degree angle facing the user or directly in front of the user. Ye et al. [56] proposed an algorithm optimizing the position and orientation of a virtual reality agent to assist users in room navigation. They modeled agent positions as quadratic functions of the distance between the user and the agent, while agent orientations were represented by polynomial functions based on the agent's position. In various contexts (i.e., school gym, museum, and factory), their approach outperformed baselines in which the agent's position was fixed relative to the object in the room (i.e., regardless of the user's position) or relative to the user (i.e., ignoring the object's position). The authors noted that automatically positioned agents could reduce the workload of human guides who relied on virtual agents for navigation assistance. Similarly, Techasarntikul et al. [49] suggested the positioning for a virtual agent in AR during a museum-guided tour to assist users in viewing the paintings. They found that users did not mind when the agent obstructed the painting but wanted them to be within the same field of view so that they did not have to trace the agent pointing to the image. Researchers also reported the effects of agent positioning on users' perceptions through monitoring physiological signals. Suzuki et al. [47] proposed a virtual agent that adjusted its position in real-time according to the arousal levels of users' heart rates. They found that the agent's irregular movements increased users' arousal levels, thereby suggesting the importance of employing a constant movement method to reach an appropriate position to suppress arousal. However, monitoring users' physiological signals for real-time agent position adjustment can be costly.

As described in this section, prior works have shown that the optimal positioning of virtual agents have a positive effect on user preference and experience in navigation tasks. As AR technology emerges to be used in shared tasks in which the user and the agent are facing the screen together, it is essential to understand how various agent positions impact users' perceptions and preferences in such context. Our approach extends prior work by placing the agent beside the user when users complete tasks to simulate a human-human interaction, like prior work noted [19], comparing to placing the agent in the user's field of view that was found to be favored by users [54]. We provide insights into user perceptions of agent positions on co-presented shared tasks.

2.3 Virtual Agent Gender

Gender has received considerable attention in diverse contexts related to virtual agents, as researchers have investigated its influence in various studies [6, 21, 27, 35]. A virtual agent's gender, whether explicitly embodied or implied through voice or behavior, can influence and impact users' preferences and interactions. For instance, Lee et al. [27] conducted a study examining human interactions with Siri, a voice-only assistant. Their findings with this intelligent virtual agent revealed that participants displayed a stronger sense of trust when their gender aligned with the gender of Siri's voice.

Other prior work has shown that users' preferences for the gender of virtual agents exhibit variability, with some results reporting participants' preferences towards female agents [15, 34, 45, 58], others for male agents [35], and some having no particular preference at all [12, 21]. Many factors can influence users' preferences regarding a virtual agent's gender like cultural norms [6] and the task at hand given gender stereotypes in society [15]. Forlizzi et al. [15] conducted a study to explore the relationship between the visual features of embodied agents and the tasks they perform. The study revealed that, in general, female agents received higher ratings compared to male agents. Additionally, the results indicated that there is a significant correlation between the task assigned to the agent and the preferred gender. This suggests that individuals tend to prefer agents whose appearance aligns with the gender stereotypes typically associated with the given task. Additionally, the authors suggested that if users do not have a clear expectation of an agent's gender for a task, the agent should just be the same gender as the user.

Krämer et al. [22] investigated the effects of the agent's rapport and gender on participants' performance, effort, and motivation on a mathematical task. The researchers found that establishing rapport with virtual agents can improve participants' performance in mathematical tasks. Contrary to their initial hypothesis, they found that virtual agents of the opposite sex are most successful at boosting participants' performance.

In summary, gender significantly impacts virtual agent design and user interactions, with varying preferences observed. While consensus on the ideal gender is lacking, aligning gender with tasks can be advantageous. Studies highlight gender's role in enhancing trust, performance, and motivation. However, most research focuses on credibility and likability. Hence, additional exploration of gender's influence on task-oriented virtual agents and user perceptions is necessary.

3 METHODOLOGY

We employed a mixed design, using agent gender as a between-subjects variable and agent representations as a within-subjects variable. To alleviate the order effect, we used a counterbalanced design for the representation of the agent and the gender of the agent. Every participant interacted with a fixed-gender agent in four different positions. This mixed design was chosen to examine both the main effects of agent gender and position, as well as their potential interaction. Prior research has shown that agent gender can significantly influence user perceptions, with users associating gendered agents with specific roles and characteristics [15, 25, 26, 43]. By examining both gender and position together, we aimed to investigate whether these factors interact to influence how users perceive virtual agents.

3.1 Participants

For this study, we recruited a total of 48 adults aged between 18 and 34 years (mean = 22, SD = 3.56 years). Users were recruited from a local university (28 self-identified as male, 20 as female).

3.2 Agents

To investigate how positioning and gender influence participants' perceptions of the agents during tasks, we presented participants with four different embodied agent configurations. Figure 1 depicts these configurations. Angel represented a human-sized agent standing beside the participant. Kai represented a human-sized agent sitting beside the participant. Nova was a small-sized agent on the table facing the participant. Drew was also a small-sized agent on the table but was oriented toward the screen.

It is important to note that certain agent configurations, such as smaller agents positioned beside the participant, either sitting or standing, or larger agents positioned atop the desk in front of the participant, were not tested in our study. This decision was influenced not only by findings from Wang et al. [54], which suggested a preference for miniature embodied agents over full-sized embodied agents, but also by the rationale that such placements would hinder effective interaction. Along these lines, agents in these configurations would either fall outside the participant's field of view, lack the ability to directly interact with the screen or participant, or represent exaggerated scenarios of users interacting with agents. These configurations were hence deemed unrealistic and were therefore excluded in our study. Each agent was available in both male and female variants, with their physical attributes remaining consistent across genders and only varying in positions. All four agents had non-verbal communication capabilities; for example, they would gaze back when the participant looked at them or make gestures toward the participant when the agent talked. The female agents each had a unique female voice, and the male agents each had a unique male voice, all sourced from Microsoft's Azure Text-to-Speech service [41]. Additionally, each agent was uniquely clothed, effectively distinguishing them from the others.

The agents were developed using the Unity editor [50] and Microsoft Mixed Reality Toolkit [37]. We used the Microsoft Rocketbox Avatar library [38] to create fully-rigged characters, and the agents were displayed on a Microsoft HoloLens 2 headset [40]. The Unity application served only as a client to display the agents and did not control the agents' dialogues, which were manipulated by the game server that ran on the same PC. The dialogue interactions were controlled using the "Wizard of Oz" technique, where predetermined responses for the virtual agent were selected in real time by the researcher pressing a designated key on a keyboard. We sent predefined hints and responses to the server, which were then sent to the HoloLens 2 to be uttered by the agent. The hints and responses for any particular puzzle were the same across all agent configurations.

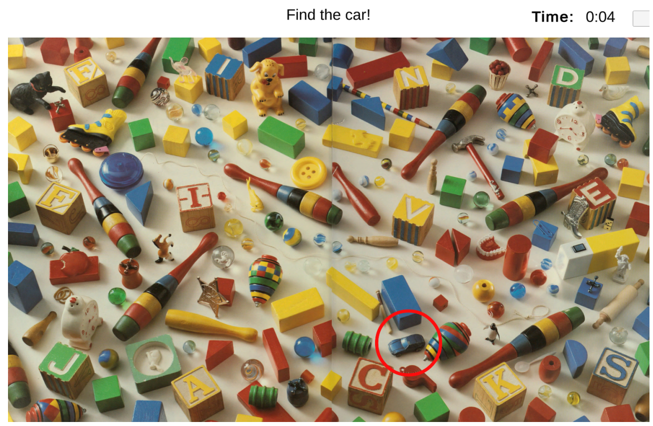

Figure 2: Illustration of the find-the-object task used in the experiment.

3.3 Hidden Object Game

We asked participants to play a game called "find-the-hidden-object". The game was inspired by Geven et al. [16] and Wang et al. [54], in which they asked people to find objects while engaging with an agent for hints. Our find-the-hidden-object game was displayed on a 23-inch/58-cm touchscreen connected to a workstation. The game was implemented in C# using the Windows Presentation Foundation framework (WPF) [39] and established a connection with the game server on the same PC, so the agents knew which puzzle the user was currently on. The task in the game was to solve puzzles from the I-Spy children's book series [31, 32]. Each puzzle is an illustration of a scene with multiple objects, including a target object hidden in the scene (as shown in Figure 2). Participants solved puzzles and tapped on the screen once they found the object.

The agents acted as assistants during the game and provided verbal hints when participants played the game. Participants were prompted to verbally ask for hints when they felt stuck. The agent could respond to utterances such as "Can you give me a hint?", and then request participants to specify the type of hint that they wanted. Notably, the agents' responses were played through the HoloLens 2 integrated speakers, with the audio spatially rendered based on the agent's location in the real world. The following types of hints were available for every puzzle:

- **object:** remind the participant of the object to find
- **color:** describes the primary color(s) of the object
- **shape:** describes the shape of the object
- **context:** give contextual cues regarding the location of the object
- **location:** describes the part of the image to find the object, such as a quadrant or half of the image

Other natural conversations were also supported, such as greetings and "you're welcome". All agents, female and male, shared the same verbal responses.

3.4 Procedures

At the beginning of the study, we obtained informed consent and then the researcher instructed the participant on how to play the game and use the HoloLens 2. Afterward, the participant wore the headset and underwent a training session to get accustomed to the game. Once the training was complete, participants were asked to finish four sets of five puzzles each, and we encouraged them to ask for hints from the agent whenever they desired. After completing each set of puzzles, participants were given a series of questionnaires. These included the NASA-TLX [18], Agent Rating Questionnaire (ARQ) [53], I-PANAS-SF [51], five questions about the agent that the participants interacted with, and one open-ended question for additional comments. Notably, the questionnaire provided after the 4th set of puzzles also incorporated demographic questions. To conclude the study, participants engaged in a post-study interview designed to gain deeper insights into their experiences and perceptions of the agents they interacted with. Specifically, we aimed to understand which agents they preferred or disliked the most and the reasons behind their preferences. Our protocol was approved by our Institutional Review Board.

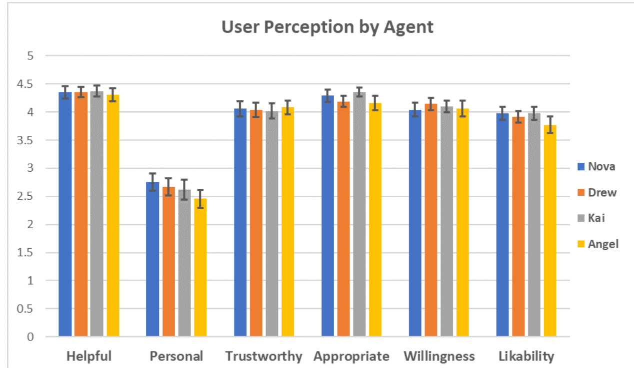

Figure 3: Average user perception metric by position.

4 RESULTS

In this section, we present the results of quantitative and qualitative analysis of participants' responses to the questionnaires for each agent, and the results of the quantitative analysis of participants' gaze data when interacting with the agents. When conducting an analysis of variance (ANOVA), we first tested for normality using a Shapiro-Wilks test. If the data distribution was not normal, we applied an Aligned Rank Transform (ART) [55] to the data before conducting a mixed-model ANOVA, which included random effects to account for repeated measures. For qualitative responses from the questionnaires, our team transformed the responses into written notes and then conducted an interpretation session of the notes. To organize and categorize the notes, we used a bottom-up approach with an affinity diagram to group related content [10]. Since most our quantitative results were not normally distributed, we initially conducted an ART-ANOVA with three independent variables: agent gender, participant gender, and agent position, aiming to explore potential influences or interaction effects of participant gender on our results. However, since we observed no interaction effect, and participant gender was a variable we did not control, we proceeded with an ART-ANOVA involving only two independent variables: agent gender and agent position. To further explore significant main effects, we conducted pairwise post-hoc comparisons using Tukey HSD corrections to control for multiple comparisons.

4.1 Agent Perceptions (ARQ)

To investigate the potential impact of virtual agent position and gender on user perceptions, we analyzed participants' responses to the ARQ questionnaire for each agent using an ART-ANOVA. Contrary to our initial expectations, the ART-ANOVA revealed no statistically significant difference in any of the six metrics of the ARQ questionnaire associated with agent position (Fig. 3). However, a significant main effect of agent gender was observed for three dimensions: *Appropriate* ($F_{1,46} = 5.29$, $p < 0.05$, $\eta^2 = 0.10$), *Willingness* ($F_{1,46} = 7.54$, $p < 0.01$, $\eta^2 = 0.14$), and *Likability* ($F_{1,46} = 8.9$, $p < 0.01$, $\eta^2 = 0.16$) (Fig. 4). Despite these differences, we did not observe any order or interaction effects between the agent position and the gender of the virtual agents.

To gain deeper insights into the gender-based preferences across the three significant scales identified in our study, we conducted pairwise post-hoc comparisons with Tukey HSD corrections. The analysis showed that, consistently, the male gender was favored the most. Specifically, participants rated male agents higher for *Likability* ($p < 0.01$), expressed greater *Willingness* to interact with male agents ($p < 0.01$), and perceived male agents as more *Appropriate* ($p < 0.05$)

While our initial analysis did not find any statistically significant results for the influence of agent position on user perceptions, the statistically significant differences observed in gender prompted us to investigate this aspect further. We divided our data into two groups: female agents and male agents. Subsequently, we conducted separate RM-ANOVA tests for each group, allowing us to delve deeper into the gender-specific differences in agent position. The analysis of male agents revealed no statistically significant difference in position on user perceptions. *Helpfulness* ($F_{3,69} = 0.45$, n.s.), *Personal* ($F_{3,69} = 0.95$, n.s.), *Trustworthiness* ($F_{3,69} = 1.01$, n.s.), *Appropriate* ($F_{3,69} = 0.1$, n.s.), *Willingness* ($F_{3,69} = 1.00$, n.s.), and *Likability* ($F_{3,69} = 0.33$, n.s.). However, it was significant in the position of the female agents on the *Personal* dimension of the ARQ questionnaire ($F_{3,69} = 2.74$, $p < 0.05$). Post-hoc tests revealed that the small-size agent facing the participants was rated more personal than the full-size side-by-side agent standing next to the participants ($p < 0.05$). We did not find significance in the *Personal* dimension for the female agent between other position settings.

4.2 Affective Responses (I-PANAS-SF)

We analyzed participants' affective responses to agent position and gender using I-PANAS-SF. Results are shown in Table 1. An ART-ANOVA revealed no statistically significant differences based on gender or position in *Positive Affect* ($F_{1,46} = 0.56$, n.s.); ($F_{3,138} = 2.19$, n.s.) or *Negative Affect* ($F_{1,46} = 0.38$, n.s.); $F_{3,138} = 0.41$, n.s.). Further analysis into each of the dimensions of the questionnaire showed no significant differences based on gender or position. Moreover, our analysis did not uncover any significant interaction effects between the agent position and the gender of the virtual agents.

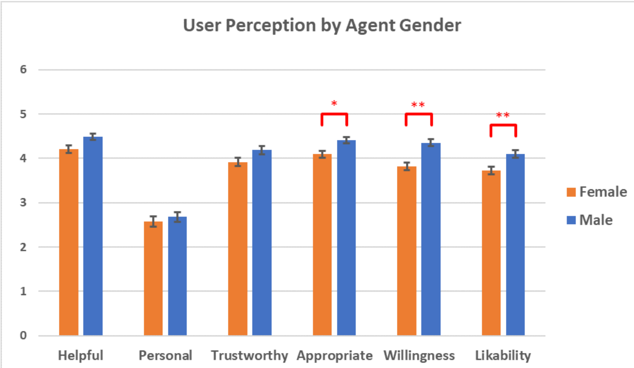

Figure 4: Average user perception metric by agent gender.

4.3 Perceived Workload (TLX)

We also explored the potential impact of virtual agents' positions and gender on participants' perceived workload while engaging in the assigned tasks. We conducted an ART-ANOVA on the results from the NASA-TLX questionnaire revealing no significant differences for either gender or position. This indicates that neither the agent's gender nor position had a discernible impact on participants' perceived workload.

4.4 Participants Subjective Preference

To examine the post-study questions regarding participants' most-liked and least-liked agents, we added up each mention as votes, similar to the method used in a study by Wang et al. [54]. Fractional votes were considered when participants could not decide on a single agent. Post-study audio data from 7 participants was not included due to technical issues with the recording equipment. All the remaining 41 valid votes were added up and divided between each of the four agents, disregarding gender differences.

In general, Nova (small desk agent facing the participant), was the most liked, with a total of 16.5 votes. Kai (human-size agent sitting beside the participant) came in second, then Drew (small desk agent facing the screen), and finally Angel (human-size agent standing beside the participant) in last place, with just 3 votes. Coincidentally, Angel was also the least-liked agent, with 25 votes, followed by Nova. Kai came in third place with Drew following closely behind.

In order to gain a more comprehensive understanding of how the aggregate preferences for the most- and least-liked options contributed to a broader consensus, we employed a net scoring approach. This involved subtracting the number of least-liked votes from the count of most-liked votes. The outcome of this analysis revealed that Kai received the highest net score of 9.5 points, closely followed by Nova with a score of 8 points. In stark contrast, Angel emerged as the most polarizing option, garnering a net score of -22, indicative of significant divergence in the participants' general subjective preferences.

To achieve a better insight into the preferences that the participants expressed in their responses, we analyzed the open-ended questions. In total, we collected 1152 responses from our 48 participants. The team discarded the blank or no comments, and then proceeded to group the responses into related clusters using a bottom-up approach with affinity diagrams, finding 13 overarching themes. The themes, which encompass comments about different agent characteristics, are: *location, size, gaze and body orientation, realness or human-likeness, gestures, appearance, interactivity and responsiveness, voice and tone, helpfulness, encouraging nature, artificial or robot-likeness, unhelpfulness*, and *pressuring nature.*

We also looked at the recurrence of comments for each theme across the different agents. All agents received virtually the same amount of positive comments regarding themes such as *helpfulness, human-likeness, appearance*, and *encouraging nature.* The agents also received similar number of negative comments in contrasting themes like *pressuring nature* and *unhelpfulness.* The following list contains the most outstanding positive and negative themes for each of our agents:

Kai (human-size agent sitting beside participant): Kai had an average number of positive comments on every theme. One participant said they liked how Female Kai was *"closer this time, more realistic"* (P28), while others said that they liked that Male Kai was *"at an eye level which felt more comforting than someone standing net [sic] to me while I'm sitting"* (P37). On the other hand, Kai agents had the most negative comments regarding *size, appearance*, and *attitude.* When talking about Male Kai, the same participant said that they *"...the life-size aspect felt a little bit odd and distracting"* (P37), while another participant remarked when commenting about Female Kai that *"when the assistants are life-size, it is easier to see their flaws"* (P38).

Nova (small desk agent facing the participant): The Nova agents had considerably more positive comments about *location* than the other agents. Observations about both Male Nova and Female Nova, such as *"It's easy to see him now"* (P45) and *"I liked having her on the desk so that she was easy to glance at while playing"* (P38), made an emphasis on why many participants liked Nova's position in the world and the fact that it was located within their field of view. However, Nova also received the most negative comments regarding its *gestures*, with comments such as *"He was shrugging and the body language didn't sit right with me"* (P45) and *"Her gestures and movements felt a little overdone."* (P38).

Drew (small desk agent facing the screen): Drew received more positive comments regarding his *attitude* and *interactivity*, while also netting just slightly more positive mentions about his *gaze direction* than the other agents. A participant said they liked that Male Drew was *"out of the way, looking at the same things that I was"* (P25), while another said they *"liked how he seemed attentive (in contrast to the previous assistants' blank stares)"* (P11). Interestingly enough, Drew also received the largest number of negative comments about their *gaze direction.* With comments such as *"she faces her body away from the user and tilts her neck to maintain constant eye contact"* (P10), and *"I did not like that he stared at me the whole time and didn't change his gaze's direction"* (P11), Drew almost doubled the number of mentions that the other agents received regarding *gaze direction* combined.

Angel (human-size agent standing beside participant): Finally, Angel, being the least preferred agent, had slightly more positive mentions than the other agents regarding their *gestures*, with participants saying they liked Angel's *"eye contact, [and] the body language"* (P43). On the other hand, *location* was the theme

Affective Response Measures to Position										
Position	Active	Alert	Attentive	Determ.	Inspired	Afraid	Ashamed	Hostile	Nervous	Upset
Nova	3.77(0.74)	3.31(1.1)	3.5(0.98)	3.46(0.89)	3.69(0.77)	1.79(0.79)	1.83(0.77)	1.67(0.8)	1.75(0.9)	1.69(0.82)
Drew	3.52(0.82)	3.04(1.08)	3.54(0.93)	3.35(0.92)	3.48(0.94)	1.67(0.8)	1.79(0.91)	1.63(0.81)	1.81(0.9)	1.77(0.82)
Kai	3.58(0.89)	3.13(1.11)	3.73(0.78)	3.5(0.87)	3.44(1.04)	1.81(0.86)	1.9(0.92)	1.63(0.73)	1.77(0.85)	1.81(0.95)
Angel	3.63(0.88)	3.17(1.09)	3.54(0.98)	3.25(1.01)	3.69(0.87)	1.88(0.83)	1.94(0.92)	1.67(0.82)	1.73(0.93)	1.69(0.85)

Affective Response Measures to Gender										
Gender	Active	Alert	Attentive	Determ.	Inspired	Afraid	Ashamed	Hostile	Nervous	Upset
M	3.53(0.83)	3.2(1.15)	3.54(0.93)	3.36(0.9)	3.59(0.85)	1.92(0.89)	1.92(0.85)	1.75(0.84)	1.72(0.91)	1.89(0.92)
F	3.72(0.84)	3.13(1.04)	3.61(0.92)	3.42(0.95)	3.55(0.98)	1.66(0.73)	1.81(0.92)	1.54(0.72)	1.81(0.88)	1.59(0.77)

Table 1: Means and standard deviations of I-PANAS-SF dimensions to both position and gender.

where Angel received both the least positive and the most negative mentions. Participants said that Angel's *location* was *"not super personable and kind of in my space"* (P25), and that Angel had a *"scary demeanor, looks down on you"* (P35). Some participants also said they felt coerced by Angel standing next to them: *"that he was standing right beside me, it makes me feel like I need to hurry"* (P21).

4.4.1 Comparing Agent Genders. We also arranged the participants' comments in agent gender groups, to better determine if major themes would arise from just comparing male versus female agents, without considering positioning or gaze direction differences. When analyzed in this manner, the positive comments for the following themes are similar between male and female agents: *size, location, gaze direction, human-likeness, gestures, appearance, helpfulness*, and *encouraging nature*. Regarding negative comments, we also observed a similar number of comments about *size, gestures*, and *interactivity*. It is important to remind the reader that we did not observe any interaction effects between agent gender and participant gender.

The next paragraphs summarize the themes that presented the largest difference in positive and negative comments count between male and female agents:

Male Agents Male agents received more positive comments about the agents' *interactivity and responsiveness*. In this regard, participants felt that male agents were quick to respond, e.g., *"he [the agent] is so accurate about what I ask and responding within no time"* (P17). Participants also made remarks about the agent being good at understanding natural speech interaction, e.g.: *"regular speech without specific commands"* (P9). On the other hand, male agents received more negative comments in themes such as *robot-likeness* and *appearance*. Participants perceived male agents as "scary" and "creepy", with comments like: *"it was creepy looking and obviously robotic sounding"* (P39), and *"I could be perceiving it this way because I already found it creepy to look at"* (P33).

Female Agents In the female agent grouping, the outstanding theme was both positive and negative comments about her *voice and tone* and how that reflected on the perceived agent's attitude. Participants' opinions regarding her *voice and tone* ranged from: *"she spoke very clearly"* (P12) to *"the virtual assistant's tone did not match the praise it was giving"* (P16). Female agents also netted the most negative comments regarding *location* and *gaze direction*, almost doubling the negative comments for male agents on this two themes. This happened even though female and male agents were placed in the same exact four locations with the same gaze direction. Finally, the *unhelpfulness* theme was also more present in female agent remarks. Participants mentioned that female agents were limited in the help they could give, with comments such as: *"I did not like how her hint system made me feel limited in what I could ask her"* (P8), or that they considered the female agent to be *"not intelligent enough"* (P44).

4.5 Gaze Interaction

In addition to analyzing our questionnaire data, we wanted to see if different agent positions and gaze directions would result in differences in gaze interactions from the participants. For all our agents, we looked at the number of times participants looked at the agent and the duration of each gaze (in seconds). Gaze estimation was based on head orientation derived from the HoloLens 2, a method similar to that used in virtual reality systems [4]. We identified gaze by checking if the participant's head orientation intersected with the designated area (referred to as a "hitbox") around the agent's body. A single gaze event was recorded when the user directed their gaze towards the agent and then shifted their gaze away.

We conducted an ART-ANOVA to analyze both gaze occurrences and gaze time per trial for agent position and agent gender. Our results revealed no statistically significant effect of gender on gaze time per trial ($F_{1,48}$ = 1.17, n.s.) or gaze occurrences per trial ($F_{1,48}$ = 0.68, n.s.). However, for position, our results revealed a statistically significant effect of agent position on both gaze occurrences per trial ($F_{3,138}$ = 7.31, $p < 0.001$) and gaze time per trial ($F_{3,138}$ = 12.06, $p < 0.001$). Post-hoc analysis with Tukey HSD corrections revealed that, for gaze occurrences, Nova, the small-size agent facing the participants, was looked at more times per trial than Drew, the small-size agent facing the screen ($p < 0.01$), Angel, the human-size agent standing beside the participants ($p < 0.05$), and Kai, the human-size agent sitting beside participants ($p < 0.001$).

5 DISCUSSION

5.1 User Perceptions on Agent Position and Gaze

The primary purpose of our research was to identify if the positioning and gaze direction of an agent in an AR setting would have any main effect on users' perceptions. While prior work has explored the placement of virtual agents in various contexts, their emphasis has been on positioning the virtual agents concerning the surrounding environment, making sure agents are not in the way of objects and are environmentally aware [11, 23, 56, 57]. In addition to placing the agent in the environment, our approach extends prior work by placing the agent beside the user to simulate a more human-like behavior. The results suggest that while positioning an agent beside a user might seem more human-like and natural, it does not necessarily lead to increased likability, reduced cognitive workload, or affect user perceptions whatsoever.

Interestingly, Nova agents (small desk agent facing the participant) garnered considerably more positive feedback regarding their location compared to other agents. Observations for both Male Nova and Female Nova, such as *"It's easy to see him now"* (P45) and *"I liked having her on the desk so that she was easy to glance at while playing"* (P38), highlighted the participants' preference for Nova's positioning, particularly appreciating its placement within their field of view. This preference also aligns with our results which revealed Nova was the most looked-at agent, with participants gazing at Nova significantly more frequently and for longer durations per trial compared to the other agents. These findings are consistent with prior work by Wang et al. [54], where participants favored smaller, desk-level agents for their novelty and reduced uncanniness compared to larger agents, and with Techasarntikul et al. [49], who revealed the users' preference for museum virtual guide agents to be within a user's immediate line of sight in museum AR setups were preferred. Nova's placement directly in front of participants ensured constant visibility and eliminated the need for users to significantly adjust their head or gaze. Additionally, Nova's front-facing orientation likely enhanced its prominence and accessibility, making it easier for participants to establish and maintain visual contact compared to the other agents. However, Nova agents received more negative comments regarding gestures than the other agents. In our implementation, gesturing was triggered only when the agents spoke under these three conditions: when users found a hidden object (praising the user), when users pressed their fingers on a wrong object (indicating that the selection was incorrect), or when users asked a question (to give more detail, to offer a hint, or to indicate that the agent was incapable of answering that particular inquiry). The gestures consisted of simple arm and hand movements, emphasizing the agent's spoken message. It is possible that Nova's positioning (facing the participant and within their field of view), made the gestures more noticeable. In contrast, both Kai and Angel were always outside of the participant's field of view, while Drew's hand movements might have been less distracting due to their arms being obscured by the agent's body angle, as shown in Figure 1.

Building on our observations, our gaze duration and frequency data suggest that all agents garnered more attention than Drew, the desk agent facing the screen. To understand this result, we need to consider that Drew was located near the participants' field of view, similar to Nova. One possible explanation for why participants' gaze was less engaging with Drew is related to the agent's body orientation and gaze direction. Our study had Drew's body facing the screen. However, we had all the agents turn their heads around and face the participants directly when observed, triggered by calculating the intersection between the hitbox and gaze. In Drew's case, this meant that if a participant pulled their head away from the screen when facing Drew, the agent would turn their head around in an almost unnatural way to look at the user directly. This phenomenon was evidenced by some responses from the open questions about Drew, such as: *"I found it a bit strange how she was looking at me over her shoulder when she could have rotated herself a bit to look at me straight on"* and *"I wish she would have faced me directly instead of looking at me over her shoulder"* (P12), and *"Creepy how it's staring at me"* (P50). It is likely that some participants were unconsciously trying to avoid looking at Drew agents, or only looked at them for the least possible amount of time, to escape from the agents' unnatural head turn and gaze.

5.2 User Perceptions on Agent Gender

The secondary purpose of our study was to identify if agent gender would have any main effect on user perceptions of virtual agents. The results of our study indicate a gender bias in the perception of virtual agents, as male agents were rated higher across dimensions of *appropriateness*, *willingness*, and *likability*. When analyzing the open-ended feedback, we can see a male gender preference in some contexts. For example, female agents received the most amount of negative comments regarding their location, gaze direction, and helpfulness, even though they were the same across all agent positions and genders.

Our findings align with prior work on the influence of gender on people's perceptions, be it human-human interactions or human-agent interactions. For example, Nass et al. [43] revealed gender stereotypical reactions to computers based on voice output. Both male and female users rated a male-voiced computer as more adept in technical subjects while rating a female-voiced computer as more knowledgeable about love and relationships. Similarly, Lee et al. [26] found that male-voiced computers had a greater influence on user decisions and were perceived as more socially attractive and trustworthy. In our study and task context, these findings could explain why participants felt the male agents were more appropriate and likable than their female counterparts, as the guidance-oriented nature of the task may have subconsciously aligned with societal stereotypes that associate masculinity with competence, authority, and problem-solving capabilities. This preference for male agents also persisted among female participants, likely reflecting the internalization of these societal norms and stereotypes regardless of their own gender [30]

Prior research has also highlighted the tendency of individuals to link gender with expertise. Typically, users align more with masculine topics when communicated by a male agent and resonate more with feminine topics when a female agent addresses them [25]. In interactions with on-screen agents, users tend to favor female agents for roles traditionally associated with women, like librarians or matchmakers, while male agents are preferred for roles often linked to men, such as athletic trainers [15]. In the context of our study, this association between gender and specific roles or tasks might provide insight into our findings. The task in our study, a "find the hidden object" game, along with the agent's role of offering hints or guidance, might subconsciously be associated with the male gender. This could be one reason why participants showed a preference for male agents in our experiment.

Given the findings of our study, combined with evidence from prior work, it is clear that gender biases and gender stereotypes can influence user's perceptions of virtual agents in human-computer interactions. These stereotypes can be understood in the context of the Social Role Theory (SRT) [13]. According to SRT, societal roles, often categorized by gender, influence our behaviors, expectations, and perceptions. As people often observe men and women in specific roles, they associate certain behaviors and attributes with each gender. Over time, these perceptions become gender stereotypes. In the context of our study, these biases and stereotypes are not just superficial, but they may improve or worsen the efficacy of virtual agents in completing certain tasks and engaging with users. While this suggests that designers should take into account gender stereotypes when designing these agents, we would argue that instead, designers should use the opposite gender roles in hopes that increasing visibility will result in breaking down these gender roles and stereotypes, similar to what researchers (e.g., [8]) advocate to reduce gender bias in fields dominated by one sex.

6 LIMITATIONS AND FUTURE WORK

In our study, we explored the effect of agent gender and position on user perceptions. However, we found some limiting factors in our study. One notable limitation was the HoloLens 2 restricted field of view, which might have led to certain agents being more consistently visible to participants than others, particularly those situated on the desk as opposed to those beside the participants. Additionally, the demographics of our participants may not represent a broader or more diverse audience, given all were recruited from a university setting. Future work could study a diverse population to uncover findings that are more widely applicable. Another limitation involves the design of our agents. Some participants felt some agents had unnatural head turns and gazes, which could have influenced their perceptions. Finally, another aspect to consider pertains to gender dynamics in our study. The chosen task could have influenced participants' perceptions regarding gender and position. Our results suggest a potential gender bias that may be linked to this specific task. Simultaneously, the study lacked control over participant gender as a variable, which could have influenced the results. Future studies should examine whether these findings remain consistent when using gender-less agents, non-gendered agents, non-human anthropomorphized agents, or androgynous agents, while also controlling for participant gender to ensure a more balanced and representative sample.

7 CONCLUSION

In this paper, we presented findings from a study examining user perceptions and interactions with virtual agents in AR, focusing on agent positioning and gender. Our data showed that while agent positioning did not significantly alter user perceptions, gender biases were evident, with male agents consistently outperforming their female counterparts in areas like appropriateness, willingness, and likability. From our qualitative analysis, Nova, the small desk agent facing the participant, emerged as the preferred agent due to its positioning, while Angel, the human-size agent standing beside participants, was the least favored, often described as intrusive, despite certain positive feedback. We also delved into how gender biases and stereotypes influenced user perceptions, even in the context of AR. These insights offer guidance and highlight the importance of addressing gender biases for the design of virtual agents in AR settings.

ACKNOWLEDGMENTS

The authors thank the reviewers of current and prior versions of this paper whose comments helped us significantly improve the paper. Special thanks to Dr. Rohith Venkatakrishnan for his invaluable feedback and insightful suggestions. We also thank the Ruiz HCI Lab for their valuable feedback. This work is partially supported by the National Science Foundation Grant Award IIS-1750840. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect these agencies' views.

REFERENCES

[1] Murat Akçayır and Gökçe Akçayır. 2017. Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational research review* 20 (2017), 1–11.

[2] 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*. 2571–2582.

[3] Sean Andrist, Tomislav Pejsa, Bilge Mutlu, and Michael Gleicher. 2012. Designing effective gaze mechanisms for virtual agents. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 705–714.

[4] Rowel Atienza, Ryan Blonna, Maria Isabel Saludares, Joel Casimiro, and Vivencio Fuentes. 2016. Interaction techniques using head gaze for virtual reality. In *2016 IEEE Region 10 Symposium (TENSYMP)*. IEEE, 110–114.

[5] Alec Azad, Jaime Ruiz, Daniel Vogel, Mark Hancock, and Edward Lank. 2012. Territoriality and behaviour on and around large vertical publicly-shared displays. In *Proceedings of the Designing Interactive Systems Conference*. 468–477.

[6] Amy Baylor, E Shen, and Xiaoxia Huang. 2003. Which pedagogical agent do learners choose? The effects of gender and ethnicity. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. Association for the Advancement of Computing in Education (AACE), 1507–1510.

[7] Christian Becker, Stefan Kopp, and Ipke Wachsmuth. 2007. Why emotions should be integrated into conversational agents. *Conversational informatics: an engineering approach* (2007), 49–68.

[8] Bettina J Casad, Jillian E Franks, Christina E Garasky, Melinda M Kittleman, Alanna C Roesler, Deidre Y Hall, and Zachary W Petzel. 2021. Gender inequality in academia: Problems and solutions for women faculty in STEM. *Journal of neuroscience research* 99, 1 (2021), 13–23.

[9] Justine Cassell, Timothy Bickmore, Mark Billinghurst, Lee Campbell, Kenny Chang, Hannes Vilhjálmsson, and Hao Yan. 1999. Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 520–527.

[10] Ji Young Cho and Eun-Hee Lee. 2014. Reducing confusion about grounded theory and qualitative content analysis: Similarities and differences. *Qualitative report* 19, 32 (2014).

[11] J.W.S. Chong, S.K. Ong, A.Y.C. Nee, and K. Youcef-Youmi. 2009. Robot programming using augmented reality: An interactive method for planning collision-free paths. *Robotics and Computer-Integrated Manufacturing* 25, 3 (2009), 689–701. https://doi.org/10.1016/j.rcim.2008.05.002

[12] Andrew J Cowell and Kay M Stanney. 2003. Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. In *International workshop on intelligent virtual agents*. Springer, 301–309.

[13] Alice H Eagly and Wendy Wood. 2012. Social role theory. *Handbook of theories of social psychology* 2 (2012), 458–476.

[14] Martin Eckert, Julia S Volmerg, Christoph M Friedrich, et al. 2019. Augmented reality in medicine: systematic and bibliographic review. *JMIR mHealth and uHealth* 7, 4 (2019), e10967.

[15] 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*. 209–221.

[16] 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* (Oslo, Norway) *(NordiCHI '06)*. Association for Computing Machinery, New York, NY, USA, 135–144. https://doi.org/10.1145/1182475.1182490

[17] Foteini Grivokostopoulou, Konstantinos Kovas, and Isidoros Perikos. 2020. The Effectiveness of Embodied Pedagogical Agents and Their Impact on Students

Learning in Virtual Worlds. *Applied Sciences* 10, 5 (2020). https://doi.org/10.3390/app10051739

[18] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, Peter A. Hancock and Najmedin Meshkati (Eds.). Human Mental Workload, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[19] Adam Kendon. 2010. *Spacing and Orientation in Co-present Interaction.* Springer Berlin Heidelberg, Berlin, Heidelberg, 1–15. https://doi.org/10.1007/978-3-642-12397-9_1

[20] Kangsoo Kim, Luke Boelling, Steffen Haesler, Jeremy Bailenson, Gerd Bruder, and Greg F. Welch. 2018. 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 *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* 105–114. https://doi.org/10.1109/ISMAR.2018.00039

[21] Tomoko Koda and Pattie Maes. 1996. Agents with faces: The effect of personification. In *Proceedings 5th IEEE international workshop on robot and human communication. RO-MAN'96 TSUKUBA.* IEEE, 189–194.

[22] Nicole C Krämer, Bilge Karacora, Gale Lucas, Morteza Dehghani, Gina Rüther, and Jonathan Gratch. 2016. Closing the gender gap in STEM with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education* 99 (2016), 1–13.

[23] Yining Lang, Wei Liang, and Lap-Fai Yu. 2019. Virtual Agent Positioning Driven by Scene Semantics in Mixed Reality. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (March 2019), 767–775. https://doi.org/10.1109/VR.2019.8798018

[24] Magic Leap. 2023. Magic Leap 2. https://www.magicleap.com/magic-leap-2 Accessed: June 8, 2023.

[25] Eun-Ju Lee. 2003. Effects of "gender" of the computer on informational social influence: the moderating role of task type. *International Journal of Human-Computer Studies* 58, 4 (2003), 347–362.

[26] Eun Ju Lee, Clifford Nass, and Scott Brave. 2000. Can computer-generated speech have gender? An experimental test of gender stereotype. In *CHI'00 extended abstracts on Human factors in computing systems.* 289–290.

[27] Sun Kyong Lee, Pavitra Kavya, and Sarah C Lasser. 2021. Social interactions and relationships with an intelligent virtual agent. *International Journal of Human-Computer Studies* 150 (2021), 102608.

[28] Tze Wei Liew and Su-Mae Tan. 2018. Exploring the effects of specialist versus generalist embodied virtual agents in a multi-product category online store. *Telematics and Informatics* 35, 1 (2018), 122–135. https://doi.org/10.1016/j.tele.2017.10.005

[29] Birgit Lugrin, Catherine Pelachaud, and David Traum. 2022. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application.* ACM.

[30] Margaret Madden. 2011. Gender stereotypes of leaders: Do they influence leadership in higher education? *Wagadu: A Journal of Transnational Women's & Gender Studies* 9, 1 (2011), 4.

[31] Jean Marzollo. 1992. I spy: a book of picture riddles. https://shop.scholastic.com/parent-ecommerce/books/i-spy-a-book-of-picture-riddles-9781338810806.html

[32] Jean Marzollo. 1993. I spy mystery: a book of picture riddles. https://shop.scholastic.com/teachers-ecommerce/teacher/books/i-spy-mystery-9780590462945.html

[33] 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 (2012), 239.

[34] 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.

[35] Helen M McBreen and Mervyn A Jack. 2001. Evaluating humanoid synthetic agents in e-retail applications. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 31, 5 (2001), 394–405.

[36] D. McNeely-White, F. Ortega, R. Beveridge, B. Draper, R. Bangar, D. Patil, J. Pustejovsky, N. Krishnaswamy, K. Rim, J. Ruiz, and I. Wang. 2019. User-Aware Shared Perception for Embodied Agents. In *2019 IEEE International Conference on Humanized Computing and Communication (HCC).* 46–51. https://doi.org/10.1109/HCC46620.2019.00015

[37] Microsoft. 2018. MixedRealityToolkit - Unity. https://github.com/Microsoft/MixedRealityToolkit-Unity

[38] Microsoft. 2022. Microsoft Rocketbox Avatar Library. https://github.com/microsoft/Microsoft-Rocketbox

[39] Microsoft. 2022. What is WPF? https://learn.microsoft.com/en-us/visualstudio/get-started/csharp/tutorial-wpf?view=vs-2022#what-is-wpf

[40] Microsoft. 2023. Microsoft Hololens 2. https://www.microsoft.com/en-us/hololens Accessed: June 8, 2023.

[41] Microsoft. 2023. Microsoft's Azure text-to-speech. https://azure.microsoft.com/en-us/products/ai-services/text-to-speech Accessed: September 13, 2023.

[42] Reza Moradinezhad and Erin T Solovey. 2021. Investigating trust in interaction with inconsistent embodied virtual agents. *International Journal of Social Robotics* 13, 8 (2021), 2103–2118.

[43] Clifford Nass, Youngme Moon, and Nancy Green. 1997. Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of applied social psychology* 27, 10 (1997), 864–876.

[44] 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.* 72–78.

[45] Andreea Niculescu, Dennis Hofs, Betsy Van Dijk, and Anton Nijholt. 2010. How the agent's gender influence users' evaluation of a QA system. In *2010 International Conference on User Science and Engineering (i-USEr).* IEEE, 16–20.

[46] Teresa K O'Leary, Elizabeth Stowell, Everlyne Kimani, Dhaval Parmar, Stefan Olafsson, Jessica Hoffman, Andrea G Parker, Michael K Paasche-Orlow, and Timothy Bickmore. 2020. Community-based cultural tailoring of virtual agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents.* 1–8.

[47] Shoudai Suzuki, Muhammad Nur Adilin Mohd Anuardi, Peeraya Sripian, Nobuto Matsuhira, and Midori Sugaya. 2020. Multi-user Robot Impression with a Virtual Agent and Features Modification According to Real-time Emotion from Physiological Signals. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN).* 1006–1012. https://doi.org/10.1109/RO-MAN47096.2020.9223585 ISSN: 1944-9437.

[48] Akikazu Takeuchi and Taketo Naito. 1995. Situated Facial Displays: Towards Social Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '95).* ACM Press/Addison-Wesley Publishing Co., USA, 450–455. https://doi.org/10.1145/223904.223965

[49] Nattaon Techasarntikul, Photchara Ratsamee, Jason Orlosky, Tomohiro Mashita, Yuki Uranishi, Kiyoshi Kiyokawa, and Haruo Takemura. 2019. Evaluation of Embodied Agent Positioning and Moving Interfaces for an AR Virtual Guide. https://doi.org/10.2312/egve.20191276

[50] Unity Technologies. 2020. Unity. https://unity.com/

[51] Edmund R. Thompson. 2007. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology* 38, 2 (2007), 227–242. https://doi.org/10.1177/0022022106297301 arXiv:https://doi.org/10.1177/0022022106297301

[52] Kristinn R Thórisson. 1999. Mind model for multimodal communicative creatures and humanoids. *Applied Artificial Intelligence* 13, 4-5 (1999), 449–486.

[53] Isaac Wang, Lea Buchweitz, Jesse Smith, Lara-Sophie Bornholdt, Jonas Grund, Jaime Ruiz, and Oliver Korn. 2020. "Wow, You Are Terrible at This!" - An Intercultural Study on Virtual Agents Giving Mixed Feedback. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (Virtual Event, Scotland, UK) *(IVA '20).* Association for Computing Machinery, New York, NY, USA, 8 pages. https://doi.org/10.1145/3383652.3423887

[54] Isaac Wang, Jesse Smith, and Jaime Ruiz. 2019. Exploring virtual agents for augmented reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* 1–12.

[55] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems.* 143–146.

[56] Zi-Ming Ye, Jun-Long Chen, Miao Wang, and Yong-Liang Yang. 2021. PAVAL: Position-Aware Virtual Agent Locomotion for Assisted Virtual Reality Navigation. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* 239–247. https://doi.org/10.1109/ISMAR52148.2021.00039 ISSN: 1554-7868.

[57] Jixuan Zhi, Lap-Fai Yu, and Jyh-Ming Lien. 2021. Designing Human-Robot Coexistence Space. *IEEE Robotics and Automation Letters* 6, 4 (2021), 7161–7168. https://doi.org/10.1109/LRA.2021.3097061

[58] John Zimmerman, Ellen Ayoob, Jodi Forlizzi, and Mick McQuaid. 2005. Putting a face on embodied interface agents. (2005).

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009