

# FVA: Modeling Perceived Friendliness of Virtual Agents Using Movement Characteristics

Tanmay Randhavane\*  
Dept of Computer Science  
University of North Carolina at Chapel Hill

Aniket Bera†  
Dept of Computer Science  
University of North Carolina at Chapel Hill

Kyra Kapsaskis‡  
Dept of Psychology and Neuroscience  
University of North Carolina at Chapel Hill

Kurt Gray§  
Dept of Psychology and Neuroscience  
University of North Carolina at Chapel Hill

Dinesh Manocha¶  
Dept of Computer Science and Electrical  
& Computer Engineering  
University of Maryland at College Park

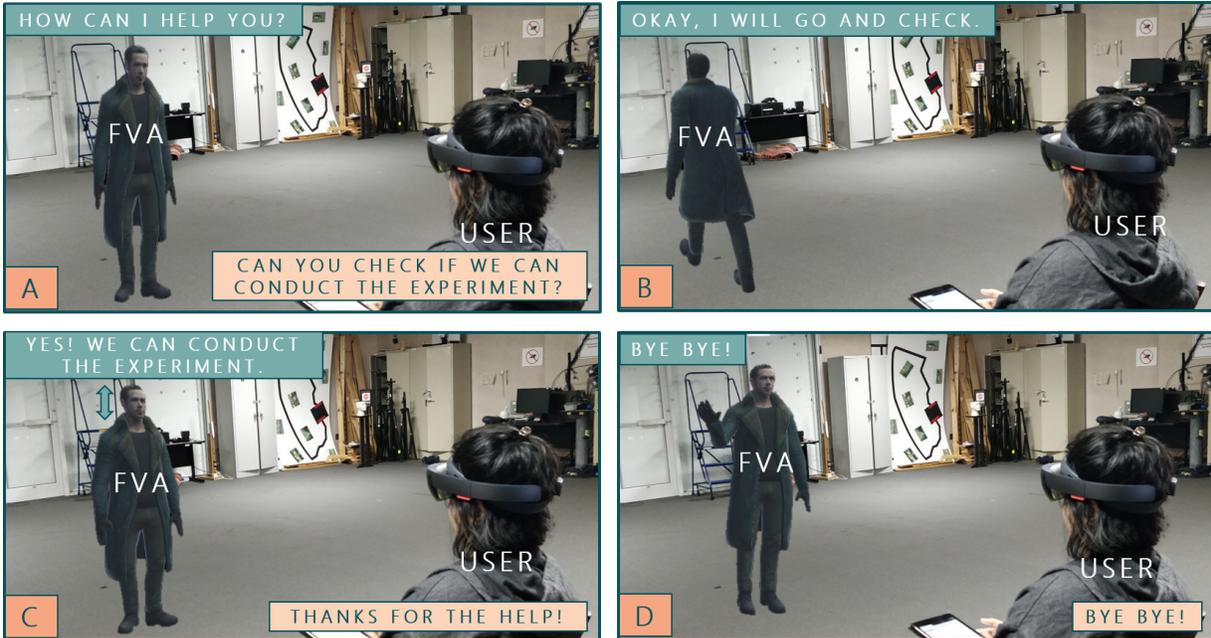


Figure 1: **Friendly Virtual Agent (FVA)**: We present an algorithm to model perceived friendliness of a virtual agent by varying its gaze (A), gait (B), and gestures corresponding to head nodding (C) and waving (D). These movement cues are generated using our novel data-driven friendliness model. We evaluate the benefits in terms of an improved sense of social presence using an AR validation study using a HoloLens.

## ABSTRACT

We present a new approach for improving the friendliness and warmth of a virtual agent in an AR environment by generating appropriate movement characteristics. Our algorithm is based on a novel data-driven friendliness model that is computed using a user-study and psychological characteristics. We use our model to control the movements corresponding to the gaits, gestures, and gazing of friendly virtual agents (FVAs) as they interact with the user’s avatar and other agents in the environment. We have integrated FVA agents with an AR environment using with a Microsoft HoloLens. Our algorithm can generate plausible movements at interactive rates to increase the social presence. We also investigate the perception of

a user in an AR setting and observe that an FVA has a statistically significant improvement in terms of the perceived friendliness and social presence of a user compared to an agent without the friendliness modeling. We observe an increment of 5.71% in the mean responses to a friendliness measure and an improvement of 4.03% in the mean responses to a social presence measure.

## 1 INTRODUCTION

Intelligent Virtual Agents (IVAs) corresponding to embodied digital or virtual characters are widely used in augmented or mixed reality environments. In many applications corresponding to treatment and rehabilitation, virtual assistance, and training, IVAs can help the user accomplish a variety of goals. Moreover, many of these applications demand that the generated IVAs look, move, behave, communicate, or act like “living” creatures, whether real or fictional.

There is considerable work on 3D modeling and generating realistic appearances of IVAs using a combination of capturing, computer vision, and rendering techniques. The resulting algorithms are being integrated with current game engines and AR systems [39, 56]. At the same time, a major issue with IVAs is that they are expected to be believable and exhibit behaviors and emotions that make them appear plausible in an AR environment. In particular, the user should

\*e-mail: tanmay@cs.unc.edu

†e-mail: ab@cs.unc.edu

‡e-mail: kykakaps@gmail.com

§e-mail: kurtjgray@gmail.com

¶e-mail: dm@cs.umd.edu

feel social presence with the IVAs and should feel confident in their abilities to perform the designated tasks [8, 31, 32]. According to prior literature [58], “social presence occurs when users feel that a form, behavior, or sensory experience indicates the presence of another individual. The amount of social presence is the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another”. It has been shown that the IVAs that exhibit social behavior and human-like personality traits can improve the sense of social presence [55].

The challenge of how to make a user trust an IVA is mirrored by an analogous issue in psychology – how to get people to trust and rely more on other people in joint tasks. When people interact with others in joint tasks, they evaluate them on two broad issues: competence and warmth/friendliness [21]. For AR applications, competence is related to an IVA’s ability and intelligence and can be increased by giving an IVA the ability to sense and affect the real world (e.g., using voice, camera, or depth sensors). Although competence is an important consideration when evaluating potential action partners, an even more important consideration in terms of evaluating others’ is warmth [11]. Warmth is one of the most fundamental interpersonal valuations and deals with whether a person is friendly or not [4]. Even if a person is competent, if they do not exhibit friendliness, then people don’t trust them and are not likely to choose them as an interaction partner. Similarly, users are likely to have more confidence in IVAs that appear friendlier.

In this paper, we address the problem of automatically generating IVA behaviors that make them appear friendlier by using appropriate cues. Humans voluntarily and involuntarily communicate emotions, moods, and intent via non-verbal cues. Friendliness or warmth is also communicated by non-verbal movement cues such as trajectories, gestures, gaits, etc. Moreover, humans use non-verbal cues such as open gestures, relaxed gaits, and moderate amounts of eye contact to communicate interpersonal warmth [54]. Inspired by these observations in psychology, we investigate methods to generate similar non-verbal movement cues to make our IVAs appear friendlier.

**Main Results:** We present an approach for modeling Friendly Virtual Agents (FVAs) in an AR environment based on a novel data-driven friendliness model. Our formulation considers three major movement characteristics: gaits, gestures, and gazing. We conduct a perception study and analyze the results to generate a mapping between gaits and perceived friendliness. Moreover, we use psychological characteristics to model the gesture and gazing features that convey friendliness. These movement characteristics are combined and we validate our *Friendliness Model* using an elaborate web-based validation study.

We use our friendliness model to generate movement characteristics in the form of non-verbal cues for an IVA. We use a Behavioral Finite State Machine (BFSM) to control the gaits, gestures, and gazing cues. Additionally, we augment these movement characteristics with global planning and collision avoidance techniques to perform navigation in the environment. We validate the impact of the generated FVA on users’ sense of social presence and their confidence in the FVA’s ability to perform designated tasks in an AR environment using a Microsoft HoloLens. Our study results suggest that FVAs provide additional confidence in their abilities to complete standard awareness and influence tasks. Additionally, users experience significantly higher social presence with the FVAs measured using the standard Temple Presence Inventory (TPI) questionnaire. The novel contributions of our work include:

1. **Friendliness Model:** We present a novel data-driven mapping between gaits and their perceived friendliness. We also present friendliness models for gestures and gazing derived using psychological characterization. We combine these three friendliness models of gaits, gestures, and gazing to form our overall Friendliness Model (Figure 2). We validate the accuracy of this model using a web-based study (Section 4).

2. We present novel algorithms to generate non-verbal movement characteristics corresponding to the gaits, gestures, and gazing using our friendliness model. Our formulation uses a BFSM, which is updated based upon the state of the environment. We use these algorithms to generate appropriate movement behaviors of IVAs in an AR environment (Section 5).
3. **AR Validation Study:** We conduct an AR study using a Microsoft HoloLens to evaluate the impact of FVAs on users’ sense of social presence and their confidence in the FVA’s ability to complete designated tasks in an immersive environment (Figure 1). We hypothesize that FVA will appear friendlier and will increase the user’s sense of presence and the confidence in the FVA’s abilities. We observe a statistically significant improvement in the perceived friendliness (5.71% improvement in mean responses) of virtual agents. Our results also suggest that users feel a statistically significant increment in the sense of social and spatial presence with FVAs (Section 6).

## 2 RELATED WORK

In this section, we give a brief survey of related work on IVAs and their use in AR and mixed reality. We also present a brief overview of movement characteristics and social behaviors.

### 2.1 Intelligent Virtual Agent Modeling

Significant research has focused on IVAs and their appearance, behaviors, and movements [45]. Norouzi et al. [50] surveyed previous work on speech modeling, verbal and non-verbal behavior, physical appearance, and identities of IVAs. Verbal approaches have been used to model different characteristics of speech and dialogues (e.g., emotions [24]). Non-verbal approaches corresponding to body posture [38], gesture [20], facial expression [42, 46], gaze [34], trajectories [9, 53], etc. have been used to model social behavior, affective behavior, and personality traits [59]. These approaches have also been used for robot navigation among pedestrians [10, 52]. Other techniques have been proposed to generate different styles of physical appearance (realistic vs. cartoon-like) and evaluate their impact on the sense of presence [25]. Other approaches reconstructed 3D avatars and poses from videos of real humans for the IVAs [39, 40, 56, 57]. Beun et al. [12] studied the impact of visual embodiment and anthropomorphism. Normoyle et al. [49] studied the effect of posture and dynamics on the perception of emotions of virtual agents. In this paper, we model IVAs with friendliness features using movement characteristics. Our approach uses gaits, gestures, and gazing features and is complimentary to these prior methods.

### 2.2 Movement Characteristics of Virtual Agents

IVAs have been used to populate AR environments for different applications [3, 51]. Majority of the research has been focused on designing virtual agents that are aware of the changes in the environment and are capable of influencing the environment [15]. Prior studies have shown that these abilities have positive effects on co-presence and social presence [37]. IVAs in AR scenarios have been evaluated on standard social presence metrics [41] and confidence indices [31]. In this paper, we design FVAs and evaluate them on similar metrics in an AR environment using a Microsoft HoloLens.

### 2.3 Social Behavior of Virtual Agents

Most previous methods have used Argyle’s model of relationship [4] for constructing virtual characters [6, 13]. According to this model, interpersonal relationships can be represented using two dimensions: dominance and friendliness. Some researchers have studied dominance relationships of IVAs [7]. In this paper, we focus on modeling the friendliness of IVAs. Friendliness refers to the closeness and

friendship levels between two or more agents. Huang et al. [30] modeled friendliness of social robots using non-verbal cues. It is known that friendliness is affected by familiarity and warmth [21]. Non-verbal cues such as gaits [43], gazing [5], facial expressions [54], etc. are used to convey, perceive and communicate warmth. Previous research has used these findings to model warmth of virtual agents. Bergmann et al. [11] used gestures to communicate warmth in virtual agents. Kim et al. [31] studied the impact of visual embodiment and social behavior on the perception of IVAs. Nguyen et al. [48] modeled the warmth and competence of virtual agents using a combination of gestures, use of space, and gaze behaviors. Our approach is inspired by these prior methods and we use non-verbal movement characteristics to generate FVAs.

### 3 OVERVIEW

In this section, we introduce our notation and the terminology used in the rest of the paper. We also give a brief overview of the psychological modeling of friendliness.

#### 3.1 Notation

Previous literature differentiates between the virtual representations perceived to be controlled by humans (i.e. avatars) or those perceived to be controlled by computer algorithms (i.e. agents). In this paper, we consider the virtual representation controlled by our algorithm as the FVA. Its visual embodiment is referred to as a 3D model, which consists of a 3D mesh along with a hierarchical skeleton. The hierarchical skeleton is represented by  $S$ , which contains the positions of all joints relative to their parent joints in a hierarchical tree with the hip joint at its root. The configuration of the skeleton at time  $t$  is represented by  $conf(S, t)$  and contains the rotation angles of all joints with respect to their parent joints.

A person’s gait or style of walking is a unique feature of their overall movement. A pose  $P \in \mathbb{R}^{48}$  of a human is a set of 3D positions of each joint  $j_i, i \in \{1, 2, \dots, 16\}$ . A gait is a set of 3D poses  $P_1, P_2, \dots, P_\tau$  where  $\tau$  is the number of time steps. Gaits and gestures correspond to time series of configurations of the skeletons. We represent a gait by  $g = [conf(S, 0), conf(S, 1), \dots, conf(S, n)]$  and a gesture by  $m = [conf(S, 0), conf(S, 1), \dots, conf(S, n)]$ . A set of gaits is represented by  $\mathbb{G}$  and a set of gestures is represented by  $\mathbb{M}$ . Gaze is an important aspect of human face-to-face interaction and can be used to increase the behavioral plausibility of the virtual characters and the overall quality of the immersive virtual experience. We begin by determining if the user is visible w.r.t the virtual agent and use that information for gaze movement. We model gazing behavior as a boolean variable that represents eye-contact represented by  $\xi$ .

We use a BFSM to trigger context dependent gestures and gaze for movement cues. The BFSM represents the state of the FVA including its immediate goals and tasks. The BFSM can also include other contexts such as mood and personality. Mathematically, we represent the BFSM by a function  $\mathbb{B} : t \times \mathbb{E} \rightarrow \mathbb{I}$ , which takes the time and state of the environment  $\mathbb{E}$  and outputs a unique BFSM state ID  $i \in \mathbb{I}$ . The environment’s state contains the positions and dimensions of the static obstacles and the current positions and velocities of the dynamic obstacles. The various movements of an IVA and its trajectory computation is performed using this BFSM.

#### 3.2 Friendliness

The notions of friendliness and warmth have been extensively studied in psychology literature. When people interact with others in performing joint tasks, they typically evaluate them on two main dimensions of social cognition: warmth and competence [19, 22]. The competence dimension concerns whether this agent is able to assist with the task at hand. This question is answered by considering various factors such as ability, intelligence, and reliability. Substantial research in psychology suggests that humans make inferences about other people based upon these factors [16, 17]. Past research

on virtual agents or human-computer interaction has also looked at these issues [11, 48].

The other dimension corresponding to the warmth or friendliness, captures whether people are friendly and well-intentioned. It deals with the most basic interpersonal judgment of whether someone is a friend or foe and whether the other person means to help or harm you. According to Fiske et al. [21], warmth is judged before competence, and it has more importance in terms of inducing affective and behavioral reactions. Since psychology literature uses warmth and friendliness in a synonymous manner, we only refer to the term friendliness in the rest of the paper.

To measure friendliness, previous studies [11] used a questionnaire of 18 items, including *pleasant, friendly, helpful* which have to be assessed on a seven-point Likert scales. Further analysis of this data revealed that the following items tend to measure friendliness with high reliability: *pleasant, sensitive, friendly, helpful, likable, approachable, sociable*. In this paper, we also use this scale to measure the friendliness and represent friendliness as a scalar value from 0 (not at all friendly) to 1 (very friendly). This scalar value is obtained by averaging the responses to the scale items and normalizing them to  $[0, 1]$ .

### 4 FRIENDLINESS MODEL

In this section, we describe the friendliness model starting with the friendliness perception study used to obtain a mapping between gaits and their perceived friendliness. Next, we describe the psychological characterizations used to model gestures and gazing features that convey friendliness. We refer to this model of friendliness and non-verbal cues as the *Friendliness Model* in this paper. Finally, we present the details of the web-based validation study used to validate our Friendliness Model.

#### 4.1 Friendliness Perception Study for Gaits

This study aimed to obtain a mapping between gaits and their perceived friendliness.

##### 4.1.1 Gait Dataset

We designed a data-driven approach and used 49 gaits from a publicly available dataset of motion-captured gaits [1] to compute this mapping. We visualized each gait in the dataset and asked the participants to rate each gait on a friendliness metric. To avoid the appearance of the 3D model affecting the gait’s perceived friendliness, we generated the gait visualizations using only a skeleton mesh (Figure 3).

##### 4.1.2 Participants

We recruited 68 participants (35 female, 33 male,  $\bar{age} = 35.75$ ) from Amazon Mechanical Turk. We presented a subset of 10 videos to each participant in a randomized order. The participant could watch each video as many times as he/she wanted and then rate the video on a friendliness measure. For each video, we obtained a minimum of 10 participant responses.

##### 4.1.3 Friendliness Measure

We used a standard friendliness measure (Section 3.2) with seven components: *pleasant, sensitive, friendly, helpful, likable, approachable, sociable*. For each component, the participants provided ratings on a 7-point scale with 7 indicating a high level of agreement. We also asked whether the participants found the gaits to be unnatural.

##### 4.1.4 Results

We first analyzed the naturalness of gaits and found that the participants found the gaits to be natural (a mean value of 3.87 for the unnaturalness question). For each gait  $g_j \in \mathbb{G}$  in the dataset, participants provided seven ratings corresponding to the items of the friendliness measure. We represent a participant response as  $r_{j,item}^k$

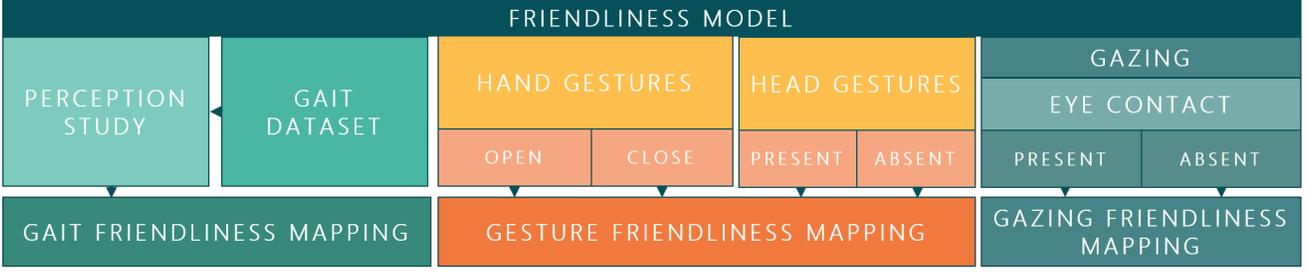


Figure 2: **Friendliness Model Overview:** Friendliness can be conveyed using nonverbal cues of gaits, gestures, and gazing. We use data-driven techniques and psychological characterization to formulate a mapping between friendliness and gaits, gestures, and gazing.

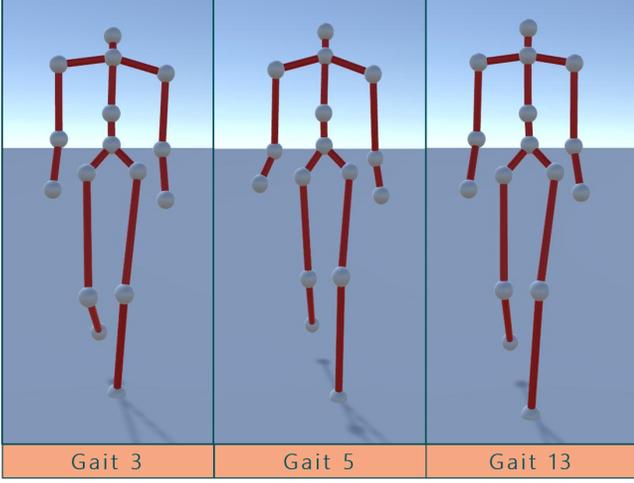


Figure 3: **Gait Visualizations:** We present sample visualizations of three gaits from a publicly-available, motion-captured gait dataset used for our user-study. We asked the participants to rate each gait video on a 7-point scale for a friendliness measure (Section 4.1.3). Based on the responses to 49 gaits from participants, we designed a data-driven model of friendliness and gaits.

where  $j$  is the gait ID,  $item$  is one of the items of the friendliness measure, and  $k$  is the participant ID. We evaluated the reliability of the measure by computing Cronbach’s  $\alpha$ . We obtained a value of  $\alpha = 0.922$  indicating that the seven components used for the friendliness measure are reliable. Since social perception is affected by the gender of the observer [14, 23, 35], we performed a t-test for differences between the responses by male and female participants. We observed that the gender of the participant did not affect the responses significantly ( $t = -0.735, p = 0.466$ ).

We aggregated the participant responses to each gait  $g_j$  and obtained the average response to each item  $r_{j,item}$  as follows:

$$r_{j,item} = \frac{\sum_{k=1}^{n_k} r_{j,item}^k}{n_k} \quad (1)$$

We compute a friendliness value  $f_j$  by aggregating the average responses to all items:

$$f_j = \frac{\sum_{item} r_{j,item}}{7} \quad (2)$$

We normalize the value of  $f_j$  to a range  $[0, 1]$  where 1 represents very high friendliness and 0 represents very low friendliness.

Using the results of this perception study, we obtained a perceived friendliness value  $f_j$  for each gait in our dataset  $g_j$ . We refer to this mapping as the friendliness model of gaits.

## 4.2 Friendliness Model for Gestures

Laban Movement Analysis [36] provides the basis for most research in the area of the description of bodily movement. According to Bergmann et al. [11], presence of gestures increases the perception of the friendliness of a virtual agent. Using video analysis, Nguyen et al. [48] divided hand gestures into two categories: open gestures and closed gestures. These results show that open gestures are perceived as more friendly, and closed gestures are perceived as less friendly. Based on these results, we connect the hand gestures  $m_{hand}$  to three levels of friendliness ( $f \in [0, 1]$ ):

$$m_{hand} = \begin{cases} \text{absent}, & \text{if } f \leq 0.33, \\ \text{closed}, & \text{if } 0.33 < f < 0.67, \\ \text{open}, & \text{if } f \geq 0.67. \end{cases} \quad (3)$$

These hand gestures include waving, folding of arms, etc. We also include appropriate head gestures (e.g., nodding and shaking of the head) for the interaction of our FVA with the user. We connect these head gestures  $m_{head}$  to high and low friendliness ( $f \in [0, 1]$ ):

$$m_{head} = \begin{cases} \text{absent}, & \text{if } f < 0.5, \\ \text{present}, & \text{if } f \geq 0.5. \end{cases} \quad (4)$$

## 4.3 Friendliness Model for Gazing

According to previous research, maintaining eye contact is associated with friendliness [48]. We control the gazing of the FVA such that the eye contact between the user and the FVA is maintained. This can be computed using visibility computations. In particular, we define the boolean variable  $\xi_f$  based on the value of friendliness as follows:

$$\xi_f = \begin{cases} 0 & \text{if } f < 0.5, \\ 1 & \text{if } f \geq 0.5, \end{cases} \quad (5)$$

where 0 represents eye-contact being absent and 1 represents eye-contact being maintained.

## 4.4 Overall Friendliness Model

We combine these three friendliness models of gaits (Equation 2), gestures (Equation 4), and gazing (Equation 5) derived using a data-driven approach and psychological characterization using a BFSM. The combined model, after combining and computing the set of gestures  $\mathbb{M}$  ( $m_{hand}$ ,  $m_{head}$ , and  $\xi_f$ ) and gait  $\mathbb{G}$  is referred to as the Friendliness Model, which provides appropriate gait, head gesture, hand gesture, and gazing behaviors according to a given friendliness level.

## 4.5 Web-based Validation Study

We conducted a web-based validation study to evaluate our model. This study aimed to validate whether our Friendliness Model was able to model the friendliness of virtual agents (FVA) correctly.

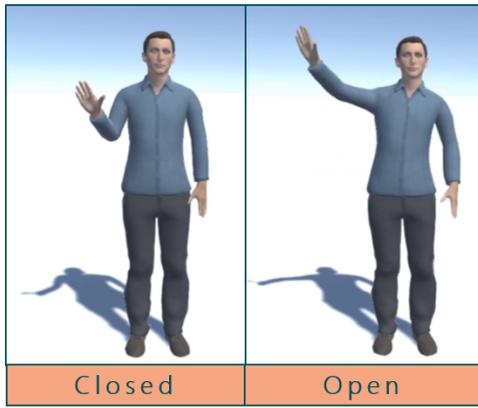


Figure 4: **Waving Gestures:** We generated videos of virtual agents with nonverbal characteristics corresponding to varying levels of friendliness as predicted by our Friendliness Model. A closed gesture corresponds to a lower friendliness level, whereas an open gesture corresponds to a higher friendliness level. We performed a web-based validation study to evaluate our model using these videos.

#### 4.5.1 Virtual Agents

We generated virtual agents with varying non-verbal cues that are modeled using our friendliness model. The virtual agents were visualized with a default male 3D model generated using Adobe Fuse software, which was rigged with a hierarchical skeleton using Adobe Mixamo. Figure 4 provides a sample visualization of the virtual agents generated for this study.

#### 4.5.2 Nonverbal Movement Characteristics

For this study, we generated videos of virtual agents with nonverbal characteristics simulated according to the following scenarios:

- **Walking:** This scenario had three videos of the virtual agent walking with three different gaits corresponding to friendliness values  $f_{low} = 0.2$ ,  $f_{medium} = 0.5$ , and  $f_{high} = 0.9$ .
- **Waving:** This scenario had three videos of the virtual agent performing a hand waving gesture with three levels of openness of the gesture.
- **Nodding:** This scenario included a video consisting of two agents standing side-by-side and answering “yes” to a question. One of the agents performed a nodding head gesture (corresponding to friendliness value  $f \geq 0.5$ ), while the other agent did not perform any head gesture (corresponding to friendliness value  $f < 0.5$ ).
- **Shaking:** This scenario included a video consisting of two agents standing side-by-side and answering “no” to a question. One of the agents performed a head shaking gesture (corresponding to friendliness value  $f \geq 0.5$ ), while the other agent did not perform any head gesture (corresponding to friendliness value  $f < 0.5$ ).
- **Gazing:** This scenario had a video consisting of two agents standing idly side-by-side. One of the agents maintained eye-contact with the user (corresponding to friendliness value  $f \geq 0.5$ ), while the other agent did not maintain eye-contact (corresponding to friendliness value  $f < 0.5$ ).

#### 4.5.3 Friendliness Measure

We again used the standard friendliness measure (Section 3.2).

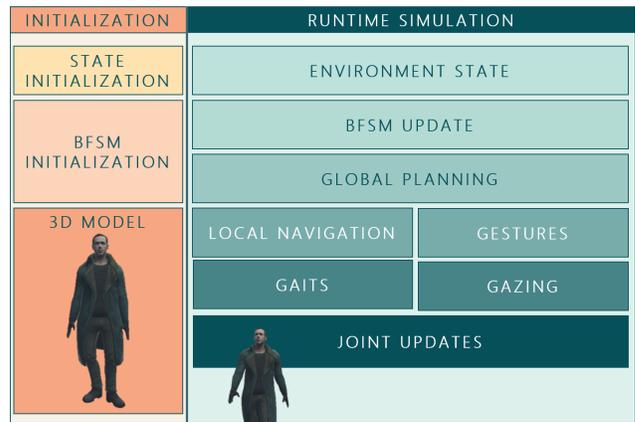


Figure 5: **FVA Generation:** We highlight our approach to generating a FVA in an AR environment. Our approach uses non-verbal movement characteristics of gaits, gestures, and gazing to convey friendliness. Based on the state-of-the-environment, these features are used to generate the movement of the FVA using a BFSM. We highlight the initialization steps on the left and various components of our runtime system on the right. The BFSM is used to compute each movement cue, which is finally used to update the various joint positions for each FVA. The FVA’s movements are visualized in an AR device (e.g., Microsoft HoloLens) using a 3D human model represented using a mesh and a hierarchical joint-based representation. We use 3D models corresponding to friends and celebrities to evaluate our approach.

#### 4.5.4 Participants

We recruited 29 participants (14 female, 15 male,  $\bar{age} = 38.82$ ) on Amazon Mechanical Turk. We presented the videos of the resulting virtual agents to participants in a randomized order. Participants rated each virtual agent on the standard friendliness measure. Therefore, for each virtual agent, we obtained 29 responses.

#### 4.5.5 Results

For each virtual agent, we aggregated the participant responses and obtained friendliness value  $f \in [0, 1]$  using the method described in Section 4.1.4. We performed a t-test for differences between the responses by male and female participants. We observed that the gender of the participant did not affect the responses significantly ( $t = -1.378, p = 0.184$ ).

We present the results of this study in Table 1. The results of this study support that our model of friendliness and nonverbal cues can model the friendliness of virtual agents.

## 5 FRIENDLY VIRTUAL AGENTS

In this section, we use our Friendliness Model to generate FVAs. We describe the algorithm to generate movements corresponding to gaits, gestures, and gazing to make an agent appear friendlier in an AR environment.

### 5.1 Movement Characteristics Generation

We provide an overview of our end-to-end approach to simulating various movements of an FVA in Figure 5. At the start of the simulation, the environment state and a BFSM are initialized based on the user input and the intended tasks. We use a 3D human model for the representation of the FVA. This model is rigged using an automatic rigging software and a hierarchical skeleton is associated with appropriate joint values.

At runtime, we update the BFSM state according to the state of the environment and whether the FVA has completed the assigned tasks. Based on the updated state, a global planning step is used to compute the next intermediate goal position and configuration for

Virtual Agent	Description	f_pred	f_obtained
Gait 1	Corresponding to $f_{low}$	$f_{low} = 0.20$	0.39
Gait 2	Corresponding to $f_{medium}$	$f_{medium} = 0.50$	0.48
Gait 3	Corresponding to $f_{high}$	$f_{high} = 0.90$	0.80
Waving 1	Closed waving	$0.33 \leq f < 0.67$	0.48
Waving 2	Slightly open waving	$f \geq 0.67$	0.72
Waving 3	Very open waving	$f \geq 0.67$	0.88
Nodding 1	No head gesture	$f < 0.5$	0.23
Nodding 2	Nods head	$f \geq 0.5$	0.89
Shaking 1	No head gesture	$f < 0.5$	0.30
Shaking 2	Shakes head	$f \geq 0.5$	0.64
Gazing 1	Maintains eye-contact	$f \geq 0.5$	0.63
Gazing 2	Does not maintain eye-contact	$f < 0.5$	0.26

Table 1: **Results of Web-based Validation Study:** We present the details of the movement characteristics generated and their friendliness values obtained using the web-based validation study. The results of this study indicate that our Friendliness Model correctly models the friendliness using nonverbal movement characteristics corresponding to gaits, gestures, and gazing.

the FVA. Next, we compute velocities for local navigation (if it needs to change location in the real world) and decide gaits, gestures, and gazing behaviors according to the BFSM state. We then update the joint positions and angles of the FVA’s hierarchical skeleton using the computed gaits, gestures, and gazing features. We visualize the FVA using the selected 3D model in an AR device such as the Microsoft HoloLens. We describe these components in detail below.

## 5.2 Behavioral Finite State Machine

We use a BFSM to represent the behavioral state of the FVA and control its behavior. We assume that the environment consists of static obstacles and dynamic obstacles. We also assume that the locations of these static obstacles and their dimensions can be reliably estimated. Additionally, dynamic obstacles can also be tracked, and their positions and velocities can be reliably determined at each time step. These dynamic obstacles may include humans in the real world (including the user) and virtual agents. During the initialization step, the environment’s state is initialized with positions and dimensions of the static obstacles and the current positions and velocities of the dynamic obstacles. We represent this state of the environment by  $\mathbb{E}$ . Using the environment state and the intended task, a BFSM is initialized and the starting state of the BFSM is computed for a given task.

At runtime, we update the state of the BFSM based on the environment state and determine the FVA’s behavior within the context of the current task. The BFSM also computes a goal position  $\vec{g}_i$  for each virtual agent. If there are obstacles or other virtual agents in the environment, we make sure that the resulting movements are collision-free.

## 5.3 Global and Local Navigation

If the goal position of virtual agents is different from their current position, then a navigation algorithm is used to compute the trajectory to the new position. We utilize the multi-agent simulation framework, *Menge* [18], to implement our local navigation algorithm. In this framework, a global navigation step breaks down the goal positions into intermediate goals that avoid collisions with the static obstacles in the environment. Moreover, a local navigation step provides collision-free navigation to the intermediate goals while avoiding collisions with dynamic obstacles, which include other virtual agents and users’ avatars, using a reciprocal collision avoidance approach. This computes a trajectory in the 2D plane. We ensure that no collisions are introduced due to other joint movements.

## 5.4 Gait Generation

When the agent is walking towards its goal position, as determined by the BFSM and the local navigation algorithm, its walking style

is modeled using its gait  $g_{des}$ . Based on the input friendliness value  $f_{des}$ , we use the Friendliness Model to obtain the gait  $g_{des}$  from the gait dataset  $\mathbb{G}$  as follows:

$$g_{des} = g_j \mid j = \min_{j=1}^n abs(f_j - f_{des}) \wedge g_j \in \mathbb{G}, \quad (6)$$

where  $n$  is the number of gaits in  $\mathbb{G}$  and  $abs(\cdot)$  represents the absolute value. This way we ensure that the gait of the FVA has the desired friendliness based on our model as presented in Section 4.1.

## 5.5 Gestures

We associate the BFSM state with gestures based on a mapping  $\mathbb{G}\mathbb{E} : \mathbb{I} \rightarrow \mathbb{M}$ , where  $\mathbb{M}$  represents a set of gestures  $m_j$  ( $j = 1 \dots n$ ) and  $\mathbb{I}$  represents the set of BFSM states. This mapping is based on the BFSM context and on the use of open gestures to convey a high level of friendliness. For example, the FVA performs a waving gesture when the FVA exits the scenario or nods its head while indicating agreement with the user. Given the appropriate hand and/or head gestures  $m_{hand}$  and  $m_{head}$ , we use the input friendliness value  $f_{des}$  and the Friendliness Model formulation (Equations 3 and 4) to decide whether to use the open or closed gesture.

## 5.6 Gazing

In addition to friendliness, maintaining eye-contact or gazing also depends on the environment state (e.g., eye-contact should not be maintained while walking away [47]). Therefore, we associate a BFSM state to decide whether to maintain eye-contact or not  $\mathbb{G}\mathbb{A} : \mathbb{I} \rightarrow \xi_{BFSM}$  where  $\xi_{BFSM}$  is a boolean variable. We combine this function with the gazing behavior decided by the Friendliness Model  $\xi_f$  (Equation 5) and compute a variable  $\xi$ :

$$\xi = \xi_f \wedge \xi_{BFSM}, \quad (7)$$

where  $\xi = 0$  represents that the eye-contact is absent and  $\xi = 1$  conveys that the eye-contact is being maintained.

We use the FVA’s 3D model to maintain eye-contact during a movement or task. Specifically, we manipulate the flexion and rotation angles of the FVA’s neck joint (Figure 6). We use the 3D positions of the user’s eye (represented by the rendering camera,  $\vec{p}_c$ ) and the position of the FVA ( $\vec{p}_{FVA}$ ) to compute the flexion  $\theta_f$  and rotation  $\theta_r$  angles:

$$\theta_f = \text{asin}\left(\frac{\vec{p}_{c,z} - \vec{p}_{FVA,z}}{\|\vec{p}_c - \vec{p}_{FVA}\|}\right), \quad (8)$$

$$\theta_r = \text{asin}\left(\frac{\vec{p}_{c,x} - \vec{p}_{FVA,x}}{\|\vec{p}_c - \vec{p}_{FVA}\|}\right). \quad (9)$$

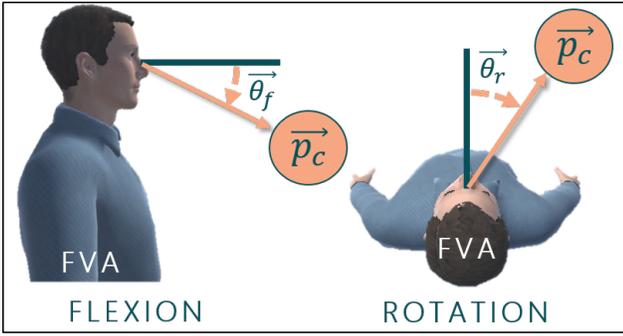


Figure 6: **Gazing Cues:** Friendliness can be conveyed using non-verbal cues associated with gazing. We model gazing features by computing the neck flexion and rotation angles such that the FVA maintains eye-contact with the user (represented by the rendering camera,  $\vec{p}_c$ ).

We can also check the line-of-sight between the agent and the user using visibility queries. For this computation, we assume that the FVA is facing the x-axis and the z-axis points vertically up. Other orientations of the FVA can be similarly modeled by adding a coordinate transform in the computation.

## 6 RESULTS AND AR VALIDATION STUDY

In this section, we describe the implementation of FVA in an AR environment using a Microsoft HoloLens. We also present the details of our validation study conducted in an AR scenario to evaluate the benefits of FVA in terms of social presence.

Our algorithm is implemented on a Windows 10 Desktop PC with an Intel Core i7-7820HK CPU at 2.90 GHz, Nvidia GeForce TitanX graphics card, and 16 GB of RAM. All the behavior modules run on the CPU, and we use the NVIDIA GPU for rendering. In practice, our system can render virtual agents at interactive rates of  $\sim 70$  FPS. Furthermore, we can handle a few tens of agents at interactive rates (i.e. at frame rates more than 30 FPS). As a user walks around the AR environment or performs the tasks, we control the friendliness of IVAs using our approach based on their gaits, gestures, and gazing. To evaluate the perceptual benefits, we also conducted an AR Validation study.

### 6.1 Experiment Goals

The goal of this experiment was to evaluate the impact of our FVAs in terms of users' sense of social presence with and their confidence in the FVA within an AR environment.

### 6.2 Participants

Based on prior studies [31] and our initial pilot tests, we decided to recruit 20 participants, which proved to be sufficient to show significant effects in our experiment. We recruited 10 female and 10 male participants for our experiment ( $age = 27.75$ ) from a university community. All the participants had a correct or corrected-to-normal vision. Three participants wore glasses during the experiment. None of the participants reported any history of discomfort with the HoloLens.

### 6.3 Virtual Agents

For this experiment, we used two types of 3D models to represent the virtual agents. One of the models, *Ryan*, was a commercially-available 3D model of a celebrity. The other model, *John*, was designed and rigged using Adobe Fuse and Adobe Mixamo. We used two different models to account for variation in appearance. Previous research suggests that females appear friendlier than males [2]. We control this gender effect by conducting the study with only male avatars. We generated virtual agents using two methods:

- **Default:** A virtual agent with a default gait (friendliness value of  $f = 0.52$ ) and with gestures and gazing features absent. We refer to this agent as the *Default* agent in the rest of this section.
- **FVA:** A virtual agent generated using our Friendliness Model. We used a gait with friendliness value of  $f = 0.97$ . The agent performed appropriate gestures, including head nodding and waving. The agent also maintained eye-contact with the participant (represented by the rendering camera).

Based on these two 3D models and two methods, we obtained four variations of virtual agents in our AR environment: *DefaultRyan*, *FVARyan*, *DefaultJohn*, *FVAJohn*.

## 6.4 Study Design

To compare the FVA with the *Default* agent, we wanted to let the participants interact with two types of agents directly. Therefore, we used a within-subjects design where each participant interacted with all four variations of the virtual agents. We presented the virtual agents in a randomized order to account for the order effects.

## 6.5 Procedure

Our experiment and evaluation were conducted in a laboratory setting. After entering the laboratory, the participants were informed about the study and the procedure. After consenting to participate in the study, the participants sat down in a chair and wore the Microsoft HoloLens HMD. We used a physical room-like experiment with an area of  $4.0m \times 4.0m$ . There was a door in front of the participant's chair that led to another room, referred to as the "adjacent room" in the rest of the paper.

Each participant performed four sessions corresponding to the four virtual agents in a randomized order. In each session, the participants and the agent interacted as part of a scripted pseudo-real story to perform some tasks based on the participant's verbal commands. This story was designed based on previously used methods to evaluate the confidence and social presence of virtual agents in an AR environment [31]. The participants were provided with instructions about how to progress through the story using an iPad. The iPad also served as an interface for answering the questionnaire for the users.

## 6.6 Interaction Scenario

The main story started with an *introduction* stage, where the virtual agent introduced itself as the "lab assistant" who will help the participant conduct an experiment. In this stage, the virtual agents *FVARyan* and *FVAJohn* maintained eye-contact with the participants, whereas *DefaultRyan* and *DefaultJohn* did not maintain eye contact. This helped us evaluate the benefits of gazing.

After the introduction stage, we begin the main story involving the participant asking the virtual agent (lab assistant) to perform some tasks in the adjacent room. Participants performed seven tasks per session related to the virtual agent's awareness of the environment (three tasks) and its influence on the environment (four tasks). We list the tasks in Table 2. The awareness tasks (A1, A2, and A3) are related to the agent's ability to sense the real world. Specifically, they involve the agent's ability to see, hear, and feel the real physical world. These abilities are analogous to humans' natural abilities to sense the world around them. For the virtual agents, these abilities can be realized with cameras, microphones, and thermometers, respectively.

The influence tasks (I1, I2, I3, and I4) are related to the virtual agent's ability to influence the real world. Specifically, they involve the agent's abilities in terms of physical, social, social critical, and digital influence. These abilities can be realized by devices such as smart-doors [44] and speakers. In our experiment, we did not implement the actual functionality and used pre-defined agent responses.

Task	Description	Participant Command	Agent Acceptance Response	Agent Completion Response
A1	See	Please check if anyone is in the adjacent room.	Okay! I am checking if anyone is in the adjacent room right now.	There are a few people in the adjacent room.
A2	Hear	Please check if it is quiet enough to perform the experiment.	Okay! I am checking if it is quiet enough to perform the experiment.	It is quiet enough to perform the experiment.
A3	Feel	Please check if the temperature is high enough to conduct the experiment.	Okay! I am checking if the temperature is high enough to conduct the experiment.	The temperature is high enough to conduct the experiment.
I1	Physical	Please close the adjacent room's other entrance.	Okay! I am closing the adjacent room's other entrance.	I closed the adjacent room's other entrance.
I2	Social	Please tell someone that the experiment will end in 15 minutes.	Okay! I am telling someone that the experiment will end in 15 minutes.	I told someone that the experiment will end in 15 minutes.
I3	Social Critical	Please tell someone that I am not feeling well.	Okay! I am telling someone that you are not feeling well.	I told someone that you are not feeling well.
I4	Digital	Please turn off the audio and video recording in the adjacent room.	Okay! I am turning off the audio and video recording in the adjacent room.	I turned off the audio and video recording in the adjacent room.

Table 2: **Tasks Performed in Our AR Validation Study:** Participants in an AR environment used a Microsoft HoloLens. Each participant was asked to give commands to the virtual agent to perform three awareness tasks and four influence tasks. The awareness tasks (A1, A2, and A3) were related to the agent's ability to sense the real world whereas the influence tasks were (I1, I2, I3, and I4) were related to the agent's ability to influence it. In each task, after the participant gave the command, the agent responded with an acceptance response and proceeded to perform the task. After completing the task, the agent provided the completion response. After each task, the participant answered a question that measured the participant's confidence in the virtual agent's ability to complete the task. After the completion of the last task, the agent responded "Bye Bye" and the simulation ended. The participants then answered a friendliness questionnaire (Section 3.2) and a subset of the Temple Presence Inventory to measure the social presence.

Within each of the seven tasks, the participant first gave a verbal command to the agent according to the instructions provided on the iPad. To avoid errors arising from speech processing, we used a human-in-the-loop mechanism where a human operator triggers the agent acceptance response after hearing the participant's commands. The agent responded to the participant via text. After providing the acceptance responses, the agent navigated to the adjacent room and performed the task. To create an impression of performing the task, the virtual agent stood facing away from the participant for five seconds. The agent then walked back and stood in front of the participant and provided a completion response. In the case of *FVARyan* and *FVAJohn*, the FVA also performed a nodding gesture while giving the completion response. After receiving the completion response, participants answered a specific task-related item using the iPad, and the agent awaited the next task. After the completion of the last task, the agent responded "Bye Bye" and the entire session of the user in the AR environment ended. In the case of *FVARyan* and *FVAJohn*, the FVA also performed a hand waving gesture accompanying the "Bye Bye" response.

## 6.7 Measures

We administered four types of measures: friendliness, awareness, influence, and social presence.

### 6.7.1 Friendliness

We again used the standard friendliness measure (Section 3.2).

### 6.7.2 Awareness and Influence

After completing each of the awareness tasks (A1, A2, and A3) and the influence tasks (I1, I2, I3, and I4), the participants rated their confidence in the abilities of the agent to complete the task. We asked questions – "How confident are you that the agent was able to ...?" with modifications according to the specific task. These questions were meant to assess the impact of our algorithms on the participants' confidence in the abilities of the agent based on the participants' immediate reactions to the virtual agents.

### 6.7.3 Social Presence

For each of the four virtual agents, after completing the seven tasks, the participants answered a subset of the Temple Presence Inventory (TPI) [41]. TPI is a questionnaire designed to assess co-presence and social presence with virtual agents. For our task, we use a slightly

modified subset of the original TPI questionnaire to measure the social presence in our AR scenario. Specifically, we considered the following subscales:

- **Social Presence:** How much does the participant feel that the virtual agent is in the same space as them and how well does the participant and the virtual agent interact/communicate?
- **Spatial presence:** How much does the participant feel that the virtual agent has come to the place where they are co-located and how much does the participant feels that they can reach out and touch the virtual agent?
- **Social richness:** How much does the participant perceive the virtual agent as sociable, warm, sensitive, personal, or intimate?
- **Engagement:** How immersive or exciting is the interaction with the virtual agent so that the participant feels deeply involved in the interaction?

## 6.8 Hypotheses

We propose the following hypotheses:

- **H1:** FVA will appear friendlier to the participants compared to the Default agent.
- **H2:** Participants will exhibit more confidence in the FVA's awareness of the real world compared to the Default agent.
- **H3:** Participants will exhibit more confidence in the FVA's ability to influence the real world compared to the Default agent.
- **H4:** Participants will feel a stronger sense of social presence with the FVA compared to the Default agent.

## 6.9 Results

We chose a familiar celebrity character (Ryan) and a default character (John) because familiarity affects friendliness [21]. Our findings corroborate these psychological findings (mean friendliness for *John* = 4.921, *Ryan* = 5.179,  $p = 0.026$ ). For further analysis, we combined participant responses to Ryan and John to show the

	Default		FVA	
	Mean	SD	Mean	SD
Friendliness	4.707	1.216	5.050	0.915
Awareness	5.075	1.239	5.308	1.082
Influence	4.938	1.149	5.056	1.136
Social Presence	4.006	1.073	4.247	1.015
Spatial Presence	3.115	0.723	3.440	1.045
Social Richness	4.346	1.232	4.729	0.969
Engagement	4.266	1.443	4.446	1.476

Table 3: **Mean Responses to Measures:** We present the participants' average responses to each of the measures for *Default* and *FVA* agents. *FVA* agents have higher means across all the measures compared to the *Default* agents.

	$\chi^2$	$p$
Friendliness	4.000	0.046
Awareness	0.030	0.862
Influence	0.118	0.732
Social Presence	4.000	0.046
Spatial Presence	6.125	0.013
Social Richness	2.189	0.139
Engagement	1.485	0.223

Table 4: **Friedman Test:** We present the  $\chi^2$  and  $p$ -values for the Friedman test performed for the *Default* vs. *FVA* comparison for different measures and subscales.

contribution of our friendliness algorithm, which is agnostic to the familiarity with the visual appearance.

We performed a t-test for differences between the responses by male and female participants. We observed that the gender of the participant did not affect the responses significantly ( $t = -0.196, p = 0.845$ ).

We present average participant responses for the four measures (including subscales) for the *Default* and the *FVA* agents in Table 3. Since the participant questionnaire responses are of an ordinal data type, we used non-parametric Friedman tests to compare between responses for the two types of virtual agents. For this test, the method used to generate the virtual agent (*Default* vs. *FVA*) is the independent variable, and the participant response is the dependent variable. We present the test statistic  $\chi^2$  and the  $p$ -value  $p$  for this test in Table 4. The results of this test reveal significant differences between the two methods for friendliness, social, and spatial presence.

## 6.10 Discussion

We now discuss these results in more details.

### 6.10.1 Friendliness

Our hypothesis H1 was concerned with the friendliness of *FVA* in an AR environment compared to the *Default* agent. We observed that there was a statistically significant difference between the two methods for the friendliness comparison ( $\chi^2 = 4.000, p = 0.046$ ) with *FVA* reporting a 5.71% higher mean compared to the *Default* agent. These results support the hypothesis H1.

### 6.10.2 Awareness and Influence

Our hypothesis H2 was concerned with the confidence in the abilities of the *FVA* to sense the real world compared to the *Default* agent. We observed that participants reported a 3.89% higher mean compared to the *Default* agent, however this difference was not significant ( $\chi^2 = 0.030, p = 0.862$ ).

Our hypothesis H3 was concerned with the confidence in the abilities of *FVA* to influence the real world compared to the *Default* agent. We observed that participants reported a 1.98% higher mean

compared to the *Default* agent, however this difference was not significant ( $\chi^2 = 0.118, p = 0.732$ ).

These results indicate that our Friendliness Model increases the users' confidence in the virtual agent's abilities to sense and influence the real world. However, the results did not significantly support hypotheses H2 and H3, possibly because these questions also attempted to evaluate the competence, which is an independent dimension of social cognition and is not affected by friendliness [21].

### 6.10.3 Social Presence

The statistically significant differences ( $\chi^2 = 4.000, p = 0.046$ ) between the two types of agents, *FVA* and *Default*, for social presence support our hypothesis H4 that the participants indeed feel a stronger sense of social presence with the *FVA* compared to the *Default* agent.

We also observed statistically significant improvement in the spatial presence ( $\chi^2 = 6.125, p = 0.013$ ). However, the improvement was not significant for social richness and engagement subscales. In optional feedback, the participants reported that they sometimes forgot that the *FVA* was a virtual agent and they wanted to wave back when the *FVA* performed the waving gesture.

## 7 CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

In this paper, we presented a novel Friendliness Model based on three non-verbal movement characteristics corresponding to gaits, gestures, and gazing. Our data-driven model was generated with a user study, and we validated its benefits using a web-based study. Based on our model, we present algorithms to interactively generate the movement characteristics of a virtual agent (*FVA*) to make it appear friendly. We performed an extensive AR validation study using Microsoft HoloLens to evaluate the benefits of our model. Our study results indicate that *FVAs* cause a statistically significant increase in the sense of social presence in an AR environment. Our study results indicate that *FVAs* cause a statistically significant increase in the sense of social and spatial presence in an AR environment (5.42% improvement in the mean participant response). Our work has both methodological and theoretical implications for psychology research, especially in terms of evaluating social psychological methods. Our study uses multiple dynamic, naturalistic channels of stimuli to reveal the integration of multiple social cues on social judgments (i.e. friendliness) and explores the basic psychological processes used by the people to perceive friendliness in others.

Our approach has some limitations. In our experiment, we did not implement the actual functionality for sensing and influencing the environment and used pre-defined agent responses. Additionally, we only consider a subset of non-verbal and movement cues to control the behavior of a virtual agent. We used a pre-determined set of gestures and a mapping between BFSM states and the gestures. Many other components also govern the perception of friendliness and warmth of virtual agents, including speech, facial expressions, and other characteristics of body language as well as cultural difference. Age and gender have been shown to affect the perception of friendliness [2, 14, 26], however in the AR validation study we did not consider these variations.

There are many avenues for future work in addition to overcoming these limitations. Learning approaches have been developed for character motion synthesis [28, 29] and speech-based automatic gesture generation [20, 27]. In the future, we want to combine our approach with these deep learning- and controller-based methods. We want to extend our approach to generate gestures with desired levels of friendliness automatically based on these learning-based approaches. Non-verbal movement cues also affect other dimensions of interpersonal trust and we would like to evaluate them. We also want to provide virtual agents with sensing and influencing capabilities using cameras, microphones, thermometers, smart-doors, speakers, etc. We want to develop an approach to designing virtual agents for AR that appear friendly as well as competent.

## ACKNOWLEDGMENTS

This research is supported in part by ARO grant W911NF-19-1-0069, Alibaba Innovative Research (AIR) program and Intel.

## REFERENCES

- [1] Cmu graphics lab motion capture database. <http://mocap.cs.cmu.edu/>, 2018.
- [2] A. E. Abele. The dynamics of masculine-agentive and feminine-communal traits: findings from a prospective study. *Journal of personality and social psychology*, 85(4):768, 2003.
- [3] J. Amores, X. Benavides, M. Comin, A. Fuste, P. Pla, and D. Miralles. Smart avatars: using avatars to interact with objects. In *Proceedings of ACM CHI*, pp. 1–3, 2012.
- [4] M. Argyle. *Bodily Communication*. 1998.
- [5] M. Argyle, L. Lefebvre, and M. Cook. The meaning of five patterns of gaze. *European journal of social psychology*, 4(2):125–136, 1974.
- [6] J. N. Bailenson, A. C. Beall, J. Blascovich, M. Raimundo, and M. Weisbuch. Intelligent agents who wear your face: Users reactions to the virtual self. In *International Workshop on Intelligent Virtual Agents*, pp. 86–99. Springer, 2001.
- [7] N. Bee, C. Pollock, E. André, and M. Walker. Bossy or wimpy: expressing social dominance by combining gaze and linguistic behaviors. In *International Conference on Intelligent Virtual Agents*, pp. 265–271. Springer, 2010.
- [8] G. Bente, S. Rüggenberg, N. C. Krämer, and F. Eschenburg. Avatar-mediated networking: Increasing social presence and interpersonal trust in net-based collaborations. *Human communication research*, 34(2):287–318, 2008.
- [9] A. Bera, T. Randhavane, E. Kubin, H. Shaik, K. Gray, and D. Manocha. Data-driven modeling of group entitativity in virtual environments. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, p. 31. ACM, 2018.
- [10] A. Bera, T. Randhavane, R. Prinja, K. Kapsaskis, A. Wang, K. Gray, and D. Manocha. The emotionally intelligent robot: Improving social navigation in crowded environments. *arXiv preprint arXiv:1903.03217*, 2019.
- [11] K. Bergmann, F. Eyssel, and S. Kopp. A second chance to make a first impression? how appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In *International Conference on Intelligent Virtual Agents*, pp. 126–138. Springer, 2012.
- [12] R.-J. Beun, E. De Vos, and C. Witteman. Embodied conversational agents: effects on memory performance and anthropomorphisation. In *International Workshop on Intelligent Virtual Agents*, pp. 315–319. Springer, 2003.
- [13] A. Cafaro, H. H. Vilhjálmsson, T. Bickmore, D. Heylen, K. R. Jóhannsdóttir, and G. S. Valgarðsson. First impressions: Users judgments of virtual agents personality and interpersonal attitude in first encounters. In *International conference on intelligent virtual agents*, pp. 67–80. Springer, 2012.
- [14] L. L. Carli, S. J. LaFleur, and C. C. Loeber. Nonverbal behavior, gender, and influence. *Journal of personality and social psychology*, 68(6):1030, 1995.
- [15] J. Chuah, A. Robb, C. White, A. Wendling, S. Lampotang, R. Kopper, and B. Lok. Exploring agent physicality and social presence for medical team training. *Presence: Teleoperators and Virtual Environments*, 22:141–170, 08 2013. doi: 10.1162/PRES\_a.00145
- [16] A. J. Cuddy, S. T. Fiske, and P. Glick. Warmth and competence as universal dimensions of social perception: The stereotype content model and the bias map. *Advances in Experimental Social Psychology*, 40:61–149, 2008.
- [17] A. J. Cuddy, P. Glick, and A. Beninger. The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in organizational behavior*, 31:73–98, 2011.
- [18] S. Curtis, A. Best, and D. Manocha. Menge: A modular framework for simulating crowd movement. *Collective Dynamics*, 1:1–40, 2016.
- [19] J. Duchan, M. Maxwell, and D. Kovarsky. Evaluating competence in the course of everyday interaction. In D. Kovarsky, J. Duchan, and M. Maxwell, eds., *Constructing (in)competence*, chap. 1, pp. 3–26. Lawrence Erlbaum, Mahwah, NJ, 1999.
- [20] Y. Ferstl and R. McDonnell. Investigating the use of recurrent motion modelling for speech gesture generation. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents, IVA 2018, Sydney, NSW, Australia, November 05-08, 2018*, pp. 93–98, 2018.
- [21] S. T. Fiske, A. J. Cuddy, and P. Glick. Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences*, 11(2):77–83, 2007.
- [22] S. T. Fiske, A. J. C. Cuddy, P. Glick, and J. Xu. *Journal of Personality and Social Psychology*, 82(6):878–902, 2002.
- [23] J. Forlizzi, J. Zimmerman, V. Mancuso, and S. Kwak. How interface agents affect interaction between humans and computers. In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, pp. 209–221. ACM, 2007.
- [24] J. Fraser, I. Papaioannou, and O. Lemon. Spoken conversational ai in video games—emotional dialogue management increases user engagement. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pp. 179–184. ACM, 2018.
- [25] M. Gerhard, D. J. Moore, and D. J. Hobbs. Continuous presence in collaborative virtual environments: Towards a hybrid avatar-agent model for user representation. In *International Workshop on Intelligent Virtual Agents*, pp. 137–155. Springer, 2001.
- [26] R. G. Harper. Power, dominance, and nonverbal behavior: An overview. In *Power, dominance, and nonverbal behavior*, pp. 29–48. Springer, 1985.
- [27] D. Hasegawa, N. Kaneko, S. Shirakawa, H. Sakuta, and K. Sumi. Evaluation of speech-to-gesture generation using bi-directional lstm network. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pp. 79–86. ACM, 2018.
- [28] D. Holden, T. Komura, and J. Saito. Phase-functioned neural networks for character control. *ACM Transactions on Graphics (TOG)*, 36(4):42, 2017.
- [29] D. Holden, J. Saito, and T. Komura. A deep learning framework for character motion synthesis and editing. *ACM Transactions on Graphics (TOG)*, 35(4):138, 2016.
- [30] C.-M. Huang, T. Iio, S. Satake, and T. Kanda. Modeling and controlling friendliness for an interactive museum robot. In *Robotics: Science and Systems*, 2014.
- [31] K. Kim, L. Boelling, S. Haesler, J. Bailenson, G. Bruder, and G. F. Welch. Does a digital assistant need a body? the influence of visual embodiment and social behavior on the perception of intelligent virtual agents in ar. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 105–114. IEEE, 2018.
- [32] K. Kim, G. Bruder, and G. Welch. Exploring the effects of observed physicality conflicts on real-virtual human interaction in augmented reality. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, p. 31. ACM, 2017.
- [33] K. Kim and G. Welch. Maintaining and enhancing human-surrogate presence in augmented reality. In *2015 IEEE International Symposium on Mixed and Augmented Reality Workshops*, pp. 15–19. IEEE, 2015.
- [34] T. Koda, T. Hirano, and T. Ishioh. Development and perception evaluation of culture-specific gaze behaviors of virtual agents. In *International Conference on Intelligent Virtual Agents*, pp. 213–222. Springer, 2017.
- [35] N. C. Krämer, B. Karacora, G. Lucas, M. Dehghani, G. Rüter, and J. Gratch. 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*, 99:1–13, 2016.
- [36] R. Laban and F. C. Lawrence. *Effort: economy of human movement*. Macdonald & Evans, 1979.
- [37] M. Lee, K. Kim, S. Daher, A. Raij, R. Schubert, J. Bailenson, and G. Welch. The wobbly table: Increased social presence via subtle incidental movement of a real-virtual table. In *2016 IEEE Virtual Reality (VR)*, pp. 11–17. IEEE, 2016.
- [38] C. Li, T. Androulakaki, A. Y. Gao, F. Yang, H. Saikia, C. Peters, and G. Skantze. Effects of posture and embodiment on social distance in human-agent interaction in mixed reality. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pp. 191–196. ACM, 2018.
- [39] H. Li. Photorealistic facial digitization and manipulation: technical perspective. *Communications of the ACM*, 62(1):95–95, 2018.
- [40] S. Lin, Y. Chen, Y.-K. Lai, R. R. Martin, and Z.-Q. Cheng. Fast capture

- of textured full-body avatar with rgb-d cameras. *The Visual Computer*, 32(6-8):681–691, 2016.
- [41] M. Lombard, T. B. Ditton, and L. Weinstein. Measuring presence: The temple presence inventory. In *In International Workshop on Presence*, p. 115, 2009.
- [42] S. Lombardi, J. Saragih, T. Simon, and Y. Sheikh. Deep appearance models for face rendering. *ACM Transactions on Graphics (TOG)*, 37(4):68, 2018.
- [43] J. Montepare and L. Zebrowitz-McArthur. Impressions of people created by age-related qualities of their gaits. *Journal of personality and social psychology*, 55:547–56, 11 1988. doi: 10.1037/0022-3514.55.4.547
- [44] M. E. Morris, B. Adair, K. Miller, E. Ozanne, R. Hansen, A. J. Pearce, N. Santamaria, L. Viega, M. Long, and C. M. Said. Smart-home technologies to assist older people to live well at home. *Journal of aging science*, 1(1):1–9, 2013.
- [45] C. Mousas, D. Anastasiou, and O. Spantidi. The effects of appearance and motion of virtual characters on emotional reactivity. *Computers in Human Behavior*, 86:99–108, 2018.
- [46] K. Nagano, J. Seo, J. Xing, L. Wei, Z. Li, S. Saito, A. Agarwal, J. Fursund, H. Li, R. Roberts, et al. pagan: real-time avatars using dynamic textures. In *SIGGRAPH Asia 2018 Technical Papers*, p. 258. ACM, 2018.
- [47] S. Narang, A. Best, T. Randhavane, A. Shapiro, and D. Manocha. Pedvr: Simulating gaze-based interactions between a real user and virtual crowds. In *Proceedings of the 22nd ACM conference on virtual reality software and technology*, pp. 91–100. ACM, 2016.
- [48] T.-H. D. Nguyen, E. Carstensdottir, N. Ngo, M. S. El-Nasr, M. Gray, D. Isaacowitz, and D. Desteno. Modeling warmth and competence in virtual characters. In *International Conference on Intelligent Virtual Agents*, pp. 167–180. Springer, 2015.
- [49] A. Normoyle, F. Liu, M. Kapadia, N. I. Badler, and S. Jörg. The effect of posture and dynamics on the perception of emotion. In *Proceedings of the ACM Symposium on Applied Perception*, pp. 91–98. ACM, 2013.
- [50] N. Norouzi, K. Kim, J. Hochreiter, M. Lee, S. Daher, G. Bruder, and G. Welch. A systematic survey of 15 years of user studies published in the intelligent virtual agents conference. In *Proceedings of the 18th international conference on intelligent virtual agents*, pp. 17–22. ACM, 2018.
- [51] G. M. O’Hare, A. G. Campbell, J. W. Stafford, and R. Aiken. Nexus: Behavioural realism in mixed reality scenarios through virtual sensing. In *18th Annual Conference on Computer Animation and Social Agents (CASA 2005), Hong Kong, China, October 17-25, 2005*, 2005.
- [52] T. Randhavane, A. Bera, E. Kubin, A. Wang, K. Gray, and D. Manocha. Pedestrian dominance modeling for socially-aware robot navigation. *2019 IEEE International Conference on Robotics and Automation (ICRA)*, 2019.
- [53] T. Randhavane, A. Bera, and D. Manocha. F2fcrowds: Planning agent movements to enable face-to-face interactions. *Presence: Teleoperators and Virtual Environments*, 26(2):228–246, 2017.
- [54] M. M. Reece and R. N. Whitman. Expressive movements, warmth, and verbal reinforcement. *The Journal of Abnormal and Social Psychology*, 64(3):234, 1962.
- [55] D. M. Romano, G. Sheppard, J. Hall, A. Miller, and Z. Ma. Basic: A believable, adaptable socially intelligent character for social presence. In *Proceedings of the 8th Annual International Workshop on Presence*, pp. 287–290. Citeseer, 2005.
- [56] P. Rubin. Facebook can make vr avatars look and move exactly like you, 2019.
- [57] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan. Scanning 3d full human bodies using kinects. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):643–650, 2012.
- [58] C. Youngblut. Experience of presence in virtual environments. Technical report, Institute For Defense Analyses, Alexandria, VA, 2003.
- [59] K. Zibrek, E. Kokkinara, and R. McDonnell. The effect of realistic appearance of virtual characters in immersive environments-does the character’s personality play a role? *IEEE transactions on visualization and computer graphics*, 24(4):1681–1690, 2018.