

The Effects of Interpersonal Attitude of a Group of Agents on User's Presence and Proxemics Behavior

ANGELO CAFARO and BRIAN RAVENET, LTCI, CNRS, Télécom ParisTech,
Université Paris-Saclay

MAGALIE OCHS, Aix Marseille Université, LSIS, DiMag

HANNES HÖGNI VILHJÁLMSOON, Reykjavik University

CATHERINE PELACHAUD, LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay

12

In the everyday world people form small conversing groups where social interaction takes place, and much of the social behavior takes place through managing interpersonal space (i.e., proxemics) and group formation, signaling their attention to others (i.e., through gaze behavior), and expressing certain attitudes, for example, friendliness, by smiling, getting close through increased engagement and intimacy, and welcoming newcomers. Many real-time interactive systems feature virtual anthropomorphic characters in order to simulate conversing groups and add plausibility and believability to the simulated environments. However, only a few have dealt with autonomous behavior generation, and in those cases, the agents' exhibited behavior should be evaluated by users in terms of appropriateness, believability, and conveyed meaning (e.g., attitudes). In this article we present an integrated intelligent interactive system for generating believable nonverbal behavior exhibited by virtual agents in small simulated group conversations. The produced behavior supports group formation management and the expression of interpersonal attitudes (friendly vs. unfriendly) both among the agents in the group (i.e., in-group attitude) and towards an approaching user in an avatar-based interaction (out-group attitude). A user study investigating the effects of these attitudes on users' social presence evaluation and proxemics behavior (with their avatar) in a three-dimensional virtual city environment is presented. We divided the study into two trials according to the task assigned to users, that is, joining a conversing group and reaching a target destination behind the group. Results showed that the out-group attitude had a major impact on social presence evaluations in both trials, whereby friendly groups were perceived as more socially rich. The user's proxemics behavior depended on both out-group and in-group attitudes expressed by the agents. Implications of these results for the design and implementation of similar intelligent interactive systems for the autonomous generation of agents' multimodal behavior are briefly discussed.

Categories and Subject Descriptors: H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Evaluation / methodology*; I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence—*Intelligent agents, Multiagent systems*; I.6.8 [Simulation and Modeling]: Types of Simulation—*Animation*; J.4 [Computer Applications]: Social and Behavioral Sciences—*Psychology; Sociology*

General Terms: Design, Experimentation, Human Factors

This work was partially supported by the EC FP7 (FP7/2007-2013) project VERVE and the H2020 European project ARIA-VALUSPA. The work has been partially conducted within the Labex SMART (ANR-11-LABX-65) "Investissements d'Avenir" program under reference ANR-11-IDEX-0004-02, supported by French state funds and managed by the ANR.

The reviewing of this article was managed by associate editor Candace Sidner.

Authors' addresses: A. Cafaro, B. Ravenet, and C. Pelachaud, Multimedia Group, LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France; emails: {angelo.cafaro, catherine.pelachaud}@telecom-paristech.fr; H. H. Vilhjálmsson, CADIA, School of Computer Science, Reykjavik University, Iceland; email: hannes@ru.is; M. Ochs (current address), Aix Marseille Université, CNRS, ENSAM, Université de Toulon, LSIS UMR7296, 13397, Marseille, France; email: magalie.ochs@lsis.org.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2016 ACM 2160-6455/2016/07-ART12 \$15.00

DOI: <http://dx.doi.org/10.1145/2914796>

Additional Key Words and Phrases: Interpersonal attitude, human territoriality, group behavior, multimodal nonverbal behavior, social presence, embodied conversational agents

ACM Reference Format:

Angelo Cafaro, Brian Ravenet, Magalie Ochs, Hannes Högni Vilhjálmsson, and Catherine Pelachaud. 2016. The effects of interpersonal attitude of a group of agents on user's presence and proxemics behavior. *ACM Trans. Interact. Intell. Syst.* 6, 2, Article 12 (July 2016), 33 pages. DOI: <http://dx.doi.org/10.1145/2914796>

1. INTRODUCTION

Many real-time interactive simulations involve the use of large open worlds populated by autonomous virtual humans¹ [Ennis et al. 2010; Szymanczyk et al. 2011]. In simulated urban environments, for example, virtual crowds of pedestrians gathering in conversing groups are deployed to improve the credibility and social richness of the scenes [Ennis and O'Sullivan 2012]. In computer games, autonomous characters (often referred as nonplayer characters or NPCs) are a powerful way to populate these worlds and strengthen the gaming experience [Lankoski 2010]. Recently, there has been a growing interest in making those characters more socially aware when players engage in an interaction with them [Bernard et al. 2008]. For example, the environments of the popular open-world videogame series *Assassin's Creed*² are populated by crowds and small groups of NPCs who fulfill daily routine activities. These characters, in addition to interacting with each other as part of the background in a scene, are becoming an integral part of the gaming experience and are expected to exhibit believable reactions when players attempt to interact with them [Szymanczyk et al. 2011].

In these environments, conversing groups can play a significant role in adding plausibility and a sense of social presence to the real-time simulation [Ennis et al. 2010]. However, often priority is not given to producing autonomously generated behaviors, thus such groups assume random formations and exhibit repeated behaviors [Ennis and O'Sullivan 2012]. One key element to increase the believability of virtual characters in a group and avoid repetitiveness of their behavior is the expression of the social attitudes between them [Lankoski 2010].

In real-life group interactions, interpersonal attitudes can be conveyed through a variety of nonverbal behaviors. For example, positive facial expressions (e.g., smiling) and increased eye contact communicate a high level of friendliness among interactants [Argyle 1988; Burgoon et al. 1984; Richmond et al. 2008]. The interpersonal space among participants may signal both adherence and cohesion to the group [Schefflen 1976; Kendon 1990], but it also indicates a stance (i.e., attitude). For example, an increasing distance between participants might indicate an overall low immediacy and a reduced friendliness [Richmond et al. 2008; Argyle 1988].

In a virtual environment, the user may approach, walk in proximity, or be a member of a group of virtual characters with his/her own avatar (i.e., player controlled character). The nonverbal expression of social attitudes, in particular friendly/unfriendly interpersonal attitudes, can be a valuable means to make the user feel immersed in the environment and perceive the characters as socially present. The characters' reaction to an approaching user can vary according to their attitudes in order to provide, for example, social inclusion with friendly behavior and even exclusion (i.e., unfriendly) when required by the scenario. Furthermore, the expression of those attitudes has the advantage of generating diversified, but still plausible, multimodal agent behavior in

¹In the remainder of the article we refer to *virtual humans* also with the following terms: *virtual characters*, *embodied conversational agents* or simply *agents*.

²<http://assassinscreed.ubi.com/en-US/home/index.aspx>.

the simulated group conversation, thus obtaining more distinct and distinguishable agents rather than flat simulations where all the agents behave similarly.

Several methods have been proposed to generate believable and diversified agent nonverbal behavior. One approach consists in exploiting motion capture data [Ennis et al. 2010; Ennis and O'Sullivan 2012]. In this article we propose a different approach. We use procedural animations, instead of motion capture data, generated in conjunction by integrating and improving two systems, Virtual Interactive Behavior (VIB) and the Impulsion Social AI Engine, for the autonomous generation of agents' multimodal behavior (both systems are described in more detail in Section 4). Impulsion originally supported the generation of nonverbal behavior for simulating a small scale group conversation (e.g., interpersonal distance handling group formation and cohesion). VIB was mainly capable of generating conversational gestures and facial expressions. The combined system autonomously generates the agents' nonverbal behavior in order to simulate a conversing group capable of expressing different interpersonal attitudes both among group members and towards the user in an interactive 3D environment.

Although the believability of a group's behavior has already been shown to increase when using the Impulsion system for handling group formations and producing supporting nonverbal behavior [Pedica et al. 2010], we have now altered some behaviors supporting group interaction (e.g., interpersonal distance) and generated others (e.g., facial expressions) to support the expression of friendly/unfriendly attitudes. The agents' behavior produced by our combined system needed a new user perception study in order to evaluate the effects on users in terms of believability and sense of social presence when interacting with the group of agents. It is important to consider that from the user's perspective, a conversing group of agents exposes both an overall group attitude expressed among the group members that we defined as *in-group attitude* (i.e., observable by the user externally as not being a member of the group), and an overall attitude expressed towards a newcomer that we defined as *out-group attitude* (i.e., directed towards the user attempting to approach and interact with the agents).

This article briefly introduces this combined system and then focuses on a study aimed at examining the effects of exhibiting nonverbal in-group and out-group interpersonal attitude on the level of social presence that users attribute to a conversing group of virtual agents in a three-dimensional (3D) virtual world. For this purpose, in two different trials, we gave users two simple tasks based on typical interaction scenarios that often occur in the real-time simulations mentioned earlier: (1) joining an ongoing conversation and (2) reaching a destination in the 3D world while passing near a group of conversing agents. Furthermore, as additional behavioral assessment we measured the user's proxemics behavior in the 3D environment (i.e., interpersonal distance of their avatar from the group of agents) while performing the tasks.

Our contributions include the following: (1) a combined intelligent system that exploits human social psychology models for the autonomous generation of an agent's nonverbal behavior supporting a group interaction and the nonverbal expression of friendly/unfriendly interpersonal attitudes; (2) an evaluation of this system aimed at studying the effects of expressed attitudes on users' sense of social presence and proxemics behavior; and (3) as a result of this evaluation, we provide insights for system designers who wish to produce real-time nonverbal behavior for plausible and expressive conversational groups in interactive 3D worlds.

The remainder of the article provides a theoretical background on expression of interpersonal attitudes and human territoriality in group interaction in Section 2. It is followed by a review of related agent work in Section 3. Section 4 briefly describes the core technologies used to conduct our experiment. The experimental design is described in Section 5. Results are shown in Section 6 and discussed in Section 7. Conclusions are presented in Section 8, and limitations and future work are discussed in Section 9.

2. THEORETICAL BACKGROUND

2.1. Nonverbal Expression of Interpersonal Attitudes

Interpersonal attitudes are essentially an individual's conscious or unconscious evaluation of how s/he feels about and relates to another person [Argyle 1988, p. 85]. Several researchers attempted to identify the dimensions that can best represent the different interpersonal attitudes expressed during social interaction. Schutz [1958] proposed the dimensions of inclusion, control, and affect. Burgoon and Hale [1984] identified 12 dimensions defining different communication styles (e.g., dominance, intimacy, affect, engagement, inclusion, and confidence). Argyle proposed a two-dimensional representation. A first dimension is the *affiliation*, characterized as the degree of liking or friendliness and ranging from unfriendly to friendly. A second dimension is the *status*, related to power and assertiveness during the interaction and ranging from submissive to dominant [Argyle 1988, p. 86].

In physical-world interactions, people express these attitudes with a great variety of nonverbal behaviors (e.g., interpersonal distance, gaze, posture, touch, and facial expressions) in both dyadic [Burgoon et al. 1984; Burgoon and Le Poire 1999] and small group interactions [Mehrabian 1969a; Cristani et al. 2011]. For instance, smiling is associated with a friendly attitude [Argyle 1988], while less-expressive facial expressions are associated with a neutral attitude [Burgoon and Le Poire 1999]. Eye contact can signal friendliness [Burgoon et al. 1984]. According to Argyle, interpersonal distance (i.e., proxemics behavior) and body orientation are one of the many ways to express unfriendly or friendly attitudes [Argyle 1988].

Interpersonal attitude can also be expressed in small group interactions. Mehrabian [1969a] describes gaze behavior, posture, and distance as important cues that impact the evaluation of attitude. For example, the amount of gaze is a parabolic function of the degree of liking towards another. It increases with the level of liking but starts decreasing when reaching the maximum value. Furthermore, a friendlier attitude is expressed assuming a closer distance following a linear function [Cristani et al. 2011].

We are using Argyle's model to represent our interpersonal attitudes. More specifically, as a first step we are focusing on the affiliation dimension (unfriendly/friendly). This model offers an intrinsic simplicity due to only two dimensions being adopted, and it has been successfully exploited in previous agents work [Lee and Marsella 2011; Ravenet et al. 2013; Chollet et al. 2014] demonstrating, for example, that judgments of a greeting agent's affiliation are quickly made in first user-agent encounters [Cafaro et al. 2012], and changes in this attitude's dimension can impact the extent to which users decide to approach and interact with virtual agents [Cafaro et al. 2013].

2.2. Conversation and Human Territoriality

When several humans gather for conversation they form groups that facilitate their interaction. While these groups may appear circular in some cases, it is important to realize that they are complex entities that adapt to the evolving social environment. Several important theories aim to clarify how conversing humans manage the space between them.

Hall [1966] described *personal space*, the space around each individual, as consisting of four concentric areas that afford different kinds of communication with other individuals. From the closest to the furthest away, Hall labeled these areas *intimate*, *personal*, *social*, and *public*. Based on his theory, most social contact between acquaintances would occur within the social area, while the personal and intimate distances are reserved for very close contact with family and good friends. This theory implies that an individual will engage in behavior to manage his/her own personal space, for example, by maintaining a certain distance to another individual to avoid unwarranted intimacy.

Schefflen [1976] focused on the management of space as a joint effort and proposed that conversations are a form of *human territorial organization* that produces certain territorial behaviors. More specifically, he identified that small group conversations have a central space called nucleus, which comprises three concentric zones starting from the innermost, a common orientational space (*o-space*), followed by a space where the participants locate (*p-space*), and surrounded by a region which is commonly used as a buffer area for potential newcomers or as a passageway for passersby (*q-space*).

Kendon [1990] suggested that participants in a conversation generally try to arrange themselves so everyone has equal, direct, and exclusive access to a common space. This implies that while the common space is shared by group members, it is defended from intrusion. Kendon calls the set of positioning and orienting behaviors that achieves this an *F-formation system*. This is a dynamic system from which the shape of the group will emerge. For example, if a new person joins the conversation, the system will ensure that a new equilibrium is found, with relatively little effort.

In Pedica and Högni Vilhjálmsón [2010], they combined a number of reactive social behaviors, including those reflecting personal space [Hall 1966], the F-formation system [Kendon 1990], and conversational zones [Schefflen 1976] in a general steering framework inspired by Reynolds [1999]. This complete management of position and orientation is the foundation of the Impulsion Social AI Engine³ that we used in the work presented in this article.

3. RELATED WORK

3.1. Expression of Interpersonal Attitudes

Several computational models for simulating agents' social attitudes through their nonverbal behavior have been proposed. For instance, Gillies and Ballin [2004] developed a general framework and a system named Demenaour based on Argyle's status and affiliation model for animating nonverbal behavior of virtual characters in improvisational visual media production and expressing interpersonal attitudes toward one another.

Lee and Marsella [2011] proposed an analysis framework of nonverbal behavior for modeling side participants and bystanders in a mixed-reality story-driven experience set in a Wild West bar. A single participant could engage in multimodal interaction with multiple virtual agents. The agents executed scripted behavior that considered interpersonal relationships, communicative acts, and conversational roles with the influence of their expressed attitudes.

Hayashi and colleagues [2012] conducted a study about human-robot encounters in public environments. In particular, a friendly attitude towards passers-by was nonverbally expressed by the robot through direct gaze towards them and by approaching a participant frontally during the encounter, whereas an unfriendly attitude was obtained by turning off gaze behavior (i.e., robot looking always ahead) and by approaching from different angles. Participants attributed to the robot expressing a friendly attitude more naturalness and a positive impression.

In Cafaro et al. [2012], they conducted a study investigating how users interpreted an agent's smile, gaze, and proxemics behavior during the first seconds of a greeting encounter in a virtual museum entrance. Users judged the agent's attitude (hostile, friendly) and personality (extraversion). Results showed that proxemics behavior accounts for an agent's extraversion, whereas a smile and gaze for accounts for friendliness.

³<http://secom.ru.is/projects/impulsion/>.

Ravenet et al. [2013] adopted a user perceptive approach for modeling the nonverbal behavior that a single agent should exhibit to convey interpersonal attitudes (both status and affiliation dimensions of Argyle's model) during face-to-face conversations with the users. They first collected a corpus using an online platform where users directly configured the behavior that they would expect from an agent for conveying different attitudes. Then they developed a Bayesian network based on the data gathered during the annotation process to create a computational model for autonomous multimodal behavior generation.

Chollet et al. [2014] proposed a computational model to express interpersonal attitude for a virtual coach by selecting sequences of nonverbal behaviors (e.g., a head nod followed by a smile).

Beck et al. [2012] studied how users interpreted the emotional body language of artificial robotic agents, including some of the nonverbal behaviors that we also consider in this article (e.g., gesture). However, they focused on the correct recognition of emotions while we study interpersonal attitudes. Furthermore, the behaviors exhibited by their agents were constrained to head movements and gestures since they used videos of NAO⁴ robots as stimuli, so they couldn't focus on facial expressions and gaze behavior as we do in this study.

In sum, these systems present certain limits. They were mainly designed for face-to-face interaction [Cafaro et al. 2012; Ravenet et al. 2013; Chollet et al. 2014], initiation of encounters [Hayashi et al. 2012], or interactive drama with multiple agents [Ballin et al. 2004; Lee and Marsella 2011]. Users interacted in a relatively fixed physical (or virtual) setting, whereas in our work we focus on a dynamic 3D environment with larger patterns of interactions, for example, dealing with several participants (e.g., a conversing group of characters) and longer locomotion distances (avatar-based user-agent interactions). Furthermore, none of these works combined the generation of behavior aimed both at regulating a multi-party interaction of a small group of characters and expressing their social attitudes.

3.2. Small Scale Dynamic Groups and Territoriality

Some computational models focused on small group simulations. For instance, Prada and Paiva [2005] modeled groups of autonomous synthetic virtual agents that collaborated with the user in the resolution of collaborative tasks within a 3D virtual environment. Their model of interaction exclusively dealt with generation of agents' nonverbal behavior oriented for a problem-solving application. Our work deals with the generation of different group nonverbal behaviors (e.g., body orientation, interpersonal distance) that abstracts from the specific application domain and supports the expression of social attitudes both among the agents in a group and during the earlier stages of user-agent interaction when users approach a group.

Rehm and Endrass [2009] implemented a toolbox for modeling the behavior of culturally influenced multi-agent systems. This tool supported the selection of different theories of social group dynamics and the rapid prototyping of produced agent nonverbal behavior influenced by the designer's chosen theory. However, this tool only considered pair-formations (user-agent), while we support bigger formations. More recently, Kistler et al. [2012] extended this work for the simulation of different cultural agent backgrounds with emphasis on gestural expressiveness and interpersonal distance behavior. They integrated the user in the virtual scenario and provided reactive agents' culture-related feedback to the user's behavior. An evaluation study of their interactive system revealed that users noticed cultural differences reflected by agents interpersonal distances. In contrast, our system focuses on the expression of social

⁴The humanoid robot produced by Aldebaran Robotics: <https://www.aldebaran.com/en>.

attitudes and considers a wider range of nonverbal behaviors. Furthermore, we support an avatar based user-agents interaction, as opposed to their full body user's tracking.

3D virtual environments that incorporate human territorial organization to produce believable and dynamic group formations tend to implement a continuous social force model [Helbing and Molnár 1995] that affects position and orientation. An early example of this is the work of Jan and Traum [2007], where the position of an individual in a group would be affected by an attractive force towards a speaker and a circular formation and repelling force from those too close and outside noise. While pioneering, their system did not include continuous control of orientation, which is an important part of territorial organization, and it is managed by the Impulsion system used in this work.

Kirchner and colleagues [2011] developed a robot capable of navigating towards a small group of people and delivering an item to a given participant. Their robot had social capabilities by adapting to the dynamic group configuration (e.g., opening) when approaching and informing the intended recipient of the delivery with gaze cues. Mutlu et al. [2012] modeled a robot's gaze behavior in three-party conversations with humans. They focused on gaze cues for establishing the roles of interlocutors (i.e., addressee or by-stander), managing the turn and signaling the discourse structure. However, these robots were not modeled for expressing social attitudes.

Jokinen et al. [2013] developed an eye gaze model for coordinating turn taking in three-party conversations. They collected casual conversational data and used an eye tracker to systematically observe participants' gaze in the interactions. Their goal was to combine eye gaze with speech cues and derive a model to predict when a participant aims to take the turn. Their work can certainly improve the realism of simulated conversing groups. In contrast, our work aims at expressing social attitudes through nonverbal cues (including gaze behavior). Moreover, we put emphasis on those behaviors that can be observed from a distance (i.e., when approaching a group), thus focusing, for example, on proxemics, gaze (excluding eye movements), and gesture behavior.

Ennis et al. [2010, 2012] first conducted a series of perceptual studies to determine the plausibility of virtual conversing groups when a limited amount of audio and motion capture data are available [2010]. Then they investigated in more detail users' sensitivity to agents' distance and orientational parameters in a simulated group conversation [2012]. In Ennis et al. [2010] their goal was first to minimize the amount of audio and motion capture data in order to simulate realistic small group conversations in urban crowds or other scenes where groups are needed (e.g., cocktail parties). Thus, they investigated users' sensitivity to visual de-synchronization (i.e., when the characters' body motions in the group are mis-aligned in time) and mis-matched audio (i.e., when characters' speech content is not matched to their gestures), and they discovered that participants were more sensitive to visual desynchronization of body motions than to mismatches between the characters' gestures and their speech content. In Ennis and O'Sullivan [2012] they found that participants were sensitive to characters' distance and body orientation to a greater extent compared to different gesturing behaviors. These two studies, as well as the study conducted by Pedica et al. [2010], suggest that taking into account human territorial behavior in group conversations is important to improve the realism of the simulation with virtual agents. In our study we consider the territoriality of participants in a group conversation, but, in addition, we modulate their nonverbal behavior (e.g., interpersonal distance) according to the attitudes that the agents aim to express both among themselves and towards the user.

3.3. User Proxemics Behavior in 3D Environments

Social behavior and norms transfer from the physical world to virtual environments (VEs) [Bailenson et al. 2003; Yee et al. 2007; Friedman et al. 2007; Nassiri et al. 2010]

and user's behavioral and physiological responses can be elicited when approached by groups of one or four virtual characters at varying distances [Llobera et al. 2010]. Interpersonal space in collaborative VEs exists and individuals maintained their personal distance when interacting with virtual humans [Nassiri et al. 2010]. The sense of presence (being in a place) and co-presence (being with other people) has been shown to improve task performance in a comparison study where a small group of humans first interacted in a 3D environment (represented by 3D avatars) and then in the real world [Slater et al. 2000].

Two observational studies in Second Life, a 3D virtual community, showed that eye contact and agents' proxemics behavior can affect the user's interpersonal distance in dyadic avatar-based interactions [Yee et al. 2007], and users when their avatars were approached by other agents populating the environment tended to exhibit distinct proxemics behavior (moving away from them) [Friedman et al. 2007].

Bailenson et al. [2003] conducted a study where participants were immersed in a 3D virtual room in which a virtual human stood. They were assigned a memory task and had to walk towards and around the agent that, in turn, was approaching them and sometimes violating their personal space. Results indicated that participants maintained greater distance from virtual humans when approaching their front compared to their back and moved farthest from virtual agents invading their space.

Immersive VEs [Bailenson et al. 2003] and an online 3D virtual community (i.e., Second Life) [Friedman et al. 2007; Yee et al. 2007] have been used to study user's proxemics behavior. However, Bailenson et al. [2003] did not deal with group of virtual agents in a nonimmersive virtual world. Furthermore, none of these studies considered the expression of interpersonal attitudes in a dynamic group.

3.4. Proxemics Behavior in Group Interaction with Robots

In Mead et al. [2013], they analyze human proxemics behavior (i.e., spatial behavior) during a social interaction involving two people and a humanoid robot. Their analysis considered individual (e.g., torso pose), physical (e.g., distance between participants), and psycho-physical (i.e., Hall's notion of interpersonal space) factors that contribute to social spacing. Results remarked the importance of considering psycho-physical factors (i.e., social norms) in real-time proxemics behavior detection when designing social robots.

Kim and colleagues [2014] studied proxemics behavior (close vs. distant) and social status (supervisor vs. subordinate) of an anthropomorphic robot in a card-matching game scenario where participants played with the robot. They found that robot's social status interacted with proxemics behavior on the overall judgments of users' experience. In particular, participants who interacted with the supervisor robot judged the experience more positive when it was close, while interactions with the subordinate robot resulted in a more positive experience when the robot was distant.

Vroon et al. [2015] focused on mobile robots' position and orientation in small group interactions. In particular, they looked at three phases of the interaction: approach, converse, and retreat. They discovered that in groups of three participants and a telepresence robot (remotely controlled), the robot can pass through a group when retreating without this action having an impact on how comfortable that retreat is for the group members.

In conclusion, we reviewed several systems that dealt with the nonverbal expression of interpersonal attitudes, but mainly for simulating face-to-face user-agent interactions. Computational models for simulating small conversing group of agents have been proposed, and perception studies demonstrated that agents' territorial behaviors (i.e., interpersonal distance and body orientation) impact the level of believability that users attribute to them, as well as the user's proxemics behavior when interacting

in a 3D environment. Proxemics behavior has been considered in human-robot interaction; however, none of the reviewed works dealt with group interpersonal attitudes. In our system we went a step further. We modulated the generated conversational and territorial behavior in a group for expressing interpersonal attitudes. Therefore, in addition to focusing on the behavior aimed at simulating a believable group of conversing agents, we tweak the produced behavior to convey an overall attitude among the group members, as well as an attitude expressed towards the user, when approaching the group during an avatar-based interaction in a 3D environment.

4. CORE SYSTEMS

In this section, first we introduce the two core systems individually. Then we describe how we combined them for simulating expressive agents' social behavior in a group and, more specifically, for creating the stimuli of our experiment presented in Section 5.

4.1. Virtual Interactive Behavior System

The VIB, formerly known as Greta, is a SAIBA-compliant (c.f. Kopp et al. 2006; Heylen et al. 2008) framework for generating virtual or robotic agent's multimodal communicative behavior in a face-to-face interaction with the user [Pecune et al. 2014]. It supports the specification and autonomous generation of an agent's communicative intents (an *Intent Planner module*). Internal configurable components support rule-based transformations of those intents into corresponding behavior plans (i.e., sequences of verbal and nonverbal behavior produced by a *Behavior Planner*) that yield animation parameters, such as movements and joint rotations represented with face and body action parameters (i.e., FAPS and BAPS).⁵ Additional components allow an agent to sense its environment (e.g., user's multimodal behavior) and to keep track of the agent's mental and emotional states.

Intent and Behavior Planners work on a single agent basis for the nonverbal expression of interpersonal attitudes. On the reception of an utterance and an intended attitude to express (e.g., assuming an agent intends to express a friendly attitude towards one or more other agents in a group conversation), supporting nonverbal behavior is produced when the agent has the turn. The ultimate outputs are joints rotations and positions for displaying various types of *conversational gestures* that vary in amplitude and speed of execution (depending on the attitude to express) and a wide range of *facial expressions* acting on the agent's face action units (AUs) [Ravenet et al. 2014].

For the overall sense of cohesion in the group interaction, in particular the management of space among agents and their F-formation, we rely on the Impulsion Social AI Engine described in the following section.

4.2. Impulsion Social AI Engine

Impulsion⁶ is a social AI engine that combines a number of reactive social behaviors, including those reflecting personal space [Hall 1966] and the F-formation system [Kendon 1990], in a general steering framework inspired by Reynolds [1999]. The engine manages important aspects including the social perception of other agents and the user (in avatar based interactions), *locomotion*, *proxemics*, and *gaze behavior*.

Its implementation relies on behavior trees and the result is a responsive and continuous steering of agents' body joints [Pedica and Vilhjálmsón 2012]. We extended Impulsion with the capability of dynamically changing agents' interpersonal space relative to other participants in a group conversation as a function of the intended

⁵See M-PEG4 standard for FAPS and BAPS here: <http://mpeg.chiariglione.org/standards/mpeg-4>.

⁶<http://secom.ru.is/projects/impulsion/>.

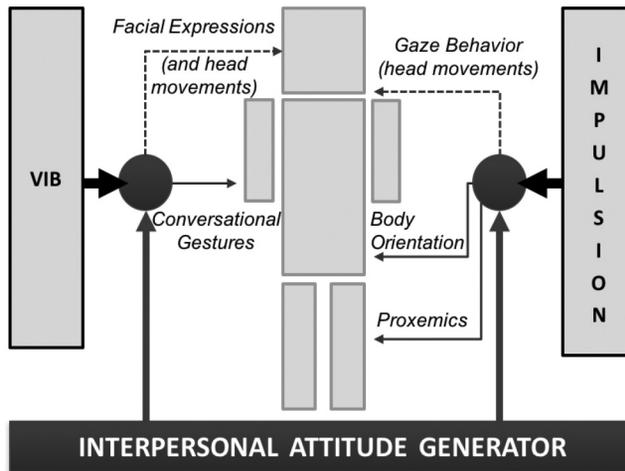


Fig. 1. An overview of the combined system architecture including *VIB* for the generation of facial expressions and conversational gestures (i.e., animation parameters), whereas *Impulsion* was used for the generation of parameters for gaze behavior (i.e., head movements), body orientation, and proxemics behavior (i.e., interpersonal distance) for the agents' group formation management. The *Interpersonal Attitude Generator* component modulates the produced parameters to reflect a friendly/unfriendly attitude in the exhibited behavior.

attitude to express, while keeping the F-formation intact. This change resulted in different irregular formations that still comply to Kendon's theory but emphasize more the interpersonal space among participants as opposed to the original implementation in which evenly distributed F-formations were displayed.

4.3. The *VIB*-*Impulsion* Combined System

VIB is a powerful framework for planning and generating agent multimodal behavior, more specifically for producing conversational gestures and facial expressions. However, its main limitation lies in the sole support of face-to-face interactions (i.e., user-agent). *Impulsion* addresses this limitation, allowing us to generate the body cues and gaze behavior for simulating a small group interaction. However, it lacks support for conversational gestures and facial expressions.

The coupling of the two systems was challenging but allowed us to obtain a powerful integrated solution to produce the agents' social behavior for expressing interpersonal attitude, for simulating small conversing groups, and, ultimately, for generating the stimuli in our experiment. The difficulty lays in the concurrent attempts to control an agent's behavior, for example, when both systems produce animation parameters for exhibiting gaze behavior. Moreover, we needed to extend *Impulsion* in order to support the parametric input of social attitudes among the agents of a group and to reflect those inputs into the interpersonal distance, body orientation, and amount of gaze behavior that participants exhibited in the simulated interaction.

We solved these issues by representing the agent's body parts as resources and deploying an overall control logic that is aware of the intended attitudes to express and handles conflicting requests for those resources. Figure 1 illustrates an overview of the integrated system design. The agent's arms and face are two resources that only *VIB* controls by generating gestural and facial expression parameters (the arrows represent these parameters indicating the affected body resources). As for the group formation, the lower body part and the torso are controlled by *Impulsion* to produce positional and orientational parameters. An *Interpersonal Attitude Generator* component (depicted in

Figure 1) modulates the produced animation parameters in order to reflect the given attitude in input (i.e., friendly vs. unfriendly) and handles conflicting requests. All components are aware of the group context and informed about the internal state of all the agents in the group. Therefore interpersonal distances, for example, are modulated according to the other agents' position. Finally, head orientation (used for exhibiting gaze behavior) is a priority-based resource (this is indicated by the dashed lines in Figure 1). During the simulated conversation, VIB can access this shared resource (i.e., when producing facial animation parameters) for exhibiting gaze behavior reflecting different attitudes of the agents in the group or conversational head movements; however, Impulsion has higher priority upon request for exhibiting a reactive gaze towards the user when in proximity of the agents.

This combined system is capable of generating the animation parameters that allowed our virtual agents to express interpersonal attitudes toward others members of the group and the user. It represents an innovative and promising technology for deploying it in future studies as well as real-time interactive environments (e.g., video games). However, some limitations still need to be overcome. First, for each group member a running instance of VIB needs to be launched, therefore preventing the system to scale up. The solution is to incorporate the control logic for handling multiple agents in a single instance of VIB and to coordinate it with Impulsion. Second, turn-taking and back-channelling are not supported yet. These are important communicative intents that frequently occur in a group conversation [Kendon 1967; Yngve 1970; Duncan 1972]. Back-channeling (i.e., active listening), for instance, can be expressed through head nods or head tilts [Bevacqua et al. 2010]. For doing this, a refined blending mechanism is needed. For example, small head movements (e.g., head nod) signalling back-channel behavior need to be blended with agents' head movements generated for exhibiting gaze behavior (e.g., head oriented towards a specific target). Due to the current limitations of the combined system, only VIB is publicly available (single executable for handling face-to-face interactions) and Impulsion is planned to be released in the near future as software library. In this article the focus is on the evaluation study presented in the next section, but we wanted to provide the reader with a general understanding of how the nonverbal behavior exhibited by our agents is produced. A more detailed description of this integration can be found in Ravenet et al. [2015].

5. EXPERIMENTAL DESIGN

We wanted to study the effects on the level of social presence that users attribute to a small conversing group. We were interested in the interplay between the social attitude that group members express both among themselves during the simulated conversation and towards the user, who is an external member of the group.

In the user-agent face-to-face scenarios reviewed in Section 3.1, the user is typically placed in front of an agent right away at the beginning of the interaction (i.e., the approaching phase is neglected) and the agent's behavior is entirely addressed to the user as explicitly defined by the social context. In our scenario, users deal with a different social context (i.e., group) that leads to interesting research questions. The members of the conversing group express an overall attitude among them through their exhibited nonverbal behavior. We define this as the *in-group attitude*. A user who is not a member of the group might experience this attitude by observing the agents behavior, but in an attempt to join them or simply walking in their proximity, the group might express an overall attitude towards him/her. We defined this as the *out-group attitude*. Therefore, when these two kind of attitudes are concurrently expressed, we had the following research questions. Which one has greater impact on user's perception of the group's social presence? Considering that the user is not



Fig. 2. The image on the left shows a technical view with the user's avatar approaching the agents and the social area of *Prudence*. She was the group member closest to the avatar and when this was within the area the nonverbal behavior supporting the out-group attitude were exhibited by group. On the right, a screen shot of the 3D world as seen by the user in third-person view with the avatar walking towards the group of agents in one of our study conditions.

immediately a member of the group, in an attempt to join it or when simply walking in proximity, do these attitudes have an influence on user's proxemics behavior?

5.1. Design Overview

5.1.1. Context. The street of a city was an open-air environment that best suited our social context. We needed to simulate a small gathering of strangers (from the user's point of view) to assign a user in the tasks described below. We aimed at addressing theoretical issues (i.e., the perception and impact of the expressed in- and out-group attitudes), but we also wanted to consider practical applications of the study outcomes. Therefore, we designed this study with the typical scenarios from gaming environments where small groups are used.

5.1.2. Tasks. We defined two tasks that considered to what extent users had to interact with the group of agents given that the two layers of attitude are expressed among the agents (i.e., in-group), indirectly perceived, and directly towards the users (i.e., out-group).

For these reasons, a first direct task in terms of user-agent interaction was to approach and join the group conversation as if the user were to ask for information about the virtual city. A second indirect task was to reach a destination in the 3D environment placed behind the conversing group of agents.

5.1.3. Apparatus and Controls. The 3D environment was displayed on a regular 27" LCD monitor and participants interacted via keyboard and mouse. They were able to move their avatar around by using the arrow keys (or the WASD keys typically adopted in videogames) and, to a limited extent, using the mouse they could look around (by pressing the left button), tilt the view up and down (right button), and zoom in and out (wheel).⁷

The users always started their task with a gender-matched virtual avatar standing far from the group. This distance (approximately 16m in the 3D world units) corresponded to the public space of the closest agent in the group (agent named *Prudence* in Figure 2 on the left) according to Hall's proxemics areas definition [Hall 1966]. This implied that while approaching the group of agents (or walking by) they could observe the ongoing simulated conversation, and at the same time the agents could react when participants were getting closer.

⁷The camera was automatically constrained in order to avoid cutting off the group from the user's viewpoint.

Table I. Summary of the Nonverbal Cues Exhibited by the Agents for All Levels of Our Independent Variables in Both Study's Trials

IV	LEVEL	DESCRIPTION
IN-GROUP	FRIENDLY	Voice medium volume, reduced speed and wide amplitude for gestures, close proximity among agents, frequent gaze at others, frequent polite smiling, long turn duration.
	NEUTRAL	Voice low volume, average speed and amplitude for gestures, medium proximity among agents, average frequency gaze at others, no smiling, neutral facial expressions and medium turn duration.
	UNFRIENDLY	Voice high volume, high speed and narrow amplitude for gestures, far proximity among agents, infrequent gaze at others, no smiling when listening, no facial expressions and short turn duration (yielding to overlaps in turn-taking).
OUT-GROUP	FRIENDLY	Smile at user's avatar, high % of gaze at user's avatar, direct body orientation towards the avatar.
	UNFRIENDLY	No smile at user's avatar, low % of gaze at user's avatar, no direct body orientation towards the avatar.

We instructed users to complete the task by pressing a key (i.e., ENTER). When the task was to join the conversation, we told them to press the key as soon as they thought they were ready to speak with the agents; however, no verbal exchange with them occurred.⁸ When the task was to reach a destination, we simply showed an arrow in the 3D space that changed color (i.e., from blue to green) when the user's avatar had reached the target point.

For each task we ran a separate trial. We named the first trial **JOIN CONVERSATION** and the second **REACH DESTINATION**. Each trial followed a within-subjects design in which participants, using their avatar, approached (in the join trial) or walked close by (in the reach trial) a series of virtual agents' groups in conversation. The stimuli only differed in the nonverbal behavior that the groups exhibited (described in the following section).

5.2. Stimuli

We manipulated the group's attitude in terms of affiliation (i.e., friendliness). In relation to the addressee to whom the attitude was expressed, we named our two independent variables (IVs) **IN-GROUP** and **OUT-GROUP**.

Table I shows a summary of the agents' behavior exhibited at all levels of our IVs in both trials. The IN-GROUP variable had three levels, respectively, *FRIENDLY*, *NEUTRAL*, and *UNFRIENDLY*. The simulated conversation among the agents had an underlying simple turn-taking mechanism that would randomly select a single speaking agent at a time and set the other group members as listeners. For each of the IN-GROUP levels, all the agents were expressing the same attitude (e.g., friendly) towards the other members of the group. Our focus was on nonverbal behavior; however, we couldn't rule out speech and accompanying gestures. Therefore, we created scripted spoken lines in French that were given in input to an English text-to-speech synthesizer. This method allowed us to keep certain audio features (e.g., voice volume) while obfuscating the verbal content of the spoken sentences. This was important for clearing out possible effects of the verbal content, which might also express certain attitudes [Argyle 1988, p. 91].

⁸Participants were allowed to stop the interaction (i.e., press ENTER) only after they had taken their avatar within the social area of the closest agent (from the avatar), see Figure 2 on the left.

The OUT-GROUP variable had two levels, as outlined in the table, respectively, *FRIENDLY* and *UNFRIENDLY*. The corresponding agents' behaviors were exhibited when the users took their avatars within the social area of the closest agent at 3.6m from it (see Figure 2 on the left, *Prudence* was the closest agent). This choice had theoretical foundations in Hall's proxemics theory [Hall 1966]. In particular, during real-world greeting interactions, in the social area, interactors use a higher level of gaze, body movements are visible, and typically smiles are exhibited [Hall 1966]. The addition of a third, neutral, level was impractical for this variable due to the kind of behavior exhibited by the agents (more details are provided below in this section). In particular, interpersonal distance and body orientation changed as a function of the group's level of friendliness (i.e., friendly groups opened to the approaching user's avatar while unfriendly groups only arranged their interpersonal space but did not orient their bodies directly towards the avatar), and a value in the middle, such as neutral, would have resulted in odd group formations.

The nonverbal behavior exhibited by the agents to produce the IN-GROUP levels was generated according to a model for expressing interpersonal attitudes described in Ravenet et al. [2015] (referred as the *Attitude Model* in the remainder of this section) and implemented in the *Interpersonal Attitude Generator* (IAG) component of our system (introduced in Section 4.3). All gestures and facial expressions of the speaking agent (i.e., the one to which the turn was randomly assigned) were timed according to the scripted spoken sentences.

The IAG generated a different voice volume (i.e., high, medium, and low) for the speaking agent with values in the continuous range [0, 1] (respectively, 1.0, 0.75, and 0.5). The voice had medium volume for the *friendly* level, high volume in the *unfriendly* level, and low volume when *neutral*. The focus here was on the relative difference among levels rather than the absolute values. These choices are based on the studies described in Argyle [1988, p. 91] (soft tone expresses friendliness, low volume is related to neutrality) and Smith-Lovin and Brody [2009, p. 408] (loud voice expresses unfriendliness).

The turn duration was short, medium, and long (respectively, 3s, 5s, and 7s). An *unfriendly* conversation yields shorter (overlapping) turns [Smith-Lovin and Brody 2009, p. 408], compared to a *neutral* one [Argyle 1988]. Longer turns appear in a *friendly* conversation [Argyle et al. 1971; Argyle 1988]. The spoken sentences were defined accordingly to be produced within these timings.

The speaking agent exhibited conversational gestures produced by *VIB*, in particular *beat* gestures according to McNeill's classification [McNeill 1992]. These gestures' speed and amplitude varied in the gesture space [McNeill 1992, p. 89] of the agent depending on the level of friendliness. The IAG modulated speed and amplitude of produced gestures according to the *Attitude Model*. This resulted in wider and slower gestures for friendlier agents, as opposed to narrower and faster gestures for unfriendly agents.

Facial expressions were blended with speech animations (i.e., lips movements). According to the *Attitude Model*, a friendly agent exhibits a polite smile more often as opposed to an unfriendly one. Thus, for a given duration of a spoken sentence to utter, the IAG produces a smile with probability $p = 0.75$ for a friendly agent, and rarely produces a smile ($p = 0.01$) for an unfriendly one. On the other hand, unfriendly agents exhibit frowns more often ($p = 0.75$) compared to friendly agents ($p = 0.11$).

The group formation is also affected by the IN-GROUP levels. In particular, the agents' interpersonal distance changed as function of the level of friendliness. The agents, arranged in the circular F-Formation system, were close to each other when the variable was at the *friendly* level, they increased the distances among them in the *neutral* condition, and they arranged with the largest distances in the *unfriendly* condition. These distances were obtained by considering the boundaries and the middle

value of an agent's personal space as defined by Hall [1966]. The personal space boundaries are then in the interval (0.45–1.30m), and the IAG attempts to assign the agents these interpersonal distances when they are, respectively, friendly (0.46m), neutral (0.87m), and unfriendly (1.28m).

The agents' gaze IN-GROUP behavior involved head movement⁹ and was exhibited for 2s on the chosen target. An agent could either look at the speaker or a desired target (another agent in the group); all the other group members had equal probability to be chosen as targets. According to the *Attitude Model*, the level of friendliness influenced the probability to have a target (as opposed to look centrally towards the group centroid) and were, respectively for all agents, at the friendly level ($p = 1$), neutral level ($p = 0.8$), and unfriendly level ($p = 0.4$).

For the OUT-GROUP variable the behavior was exhibited by all agents in a continuous and synchronized manner. While each group member was constantly adapting its positioning and orienting behaviors towards the other members for holding the F-formation system (i.e., group conversational setting), at the same time all group members exhibited the nonverbal cues towards the avatar according to the (decreasing) distance between the avatar and the group. A first set of cues (i.e., smile and gaze behavior) started to be triggered when the avatar was within the social area of *Prudence* (see Figure 2 on the left), and the group arrangement changed when the avatar walked to the end of the social area (i.e., close to the group). More specifically, in the *friendly* condition the agents smiled when the avatar entered the social area. Furthermore, the "high %" gaze was obtained with a 2s eye glance at the avatar followed by another 2s glance when the first glance ended and the group opened to make space for the user's avatar (i.e., the agents oriented their bodies towards the avatar and re-arranged their interpersonal distances). In the *unfriendly* condition the agents did not smile, but they glanced for 2s when the avatar entered the social area (only once). When the avatar reached the end of the social area, the agents kept arranging their interpersonal distance considering the presence of the avatar but they did not directly orient their bodies towards it. This made it possible to keep active the group cohesion, showing to the user that the group noticed the avatar, but at the same time expressed an unfriendly (noninclusive) attitude. In both cases, they restarted the conversation afterwards.

5.2.1. Manipulation Check. We first conducted a manipulation check (N = 12, 5 females and 7 males) to ensure that the agents' attitudes (i.e., stimuli of the main study) were correctly perceived. For this reason, we focused on a single level of the IVs at time. For the IN-GROUP check participants observed the three conditions through a fixed camera, thus excluding the possible effects of OUT-GROUP. For the OUT-GROUP check, participants used their avatar the same way as in the main experiment and we set the in-group level to neutral while varying the two out-group levels (i.e., friendly vs. unfriendly). After each of these checks, participants reported their perceived affiliation of the group on a 5-point Likert scale ranging from *very unfriendly* to *very friendly* (outcomes in the continuous interval 0–4). We also asked them to provide general feedback. After two iterations we fine tuned the values for the turn duration and voice volume of the agents as reported earlier, and we also refined the naturalness and believability of produced gesture and facial animations. For IN-GROUP, the average scores of friendliness that participants attributed to the three levels reflected the intended attitude expressed by our system, respectively, for *unfriendly* level ($M = 0.75$, $S.E. = .7$), *neutral* ($M = 2.25$, $S.E. = .5$), and *friendly* ($M = 2.67$, $S.E. = .2$). Similarly, for the OUT-GROUP variable, subject scores were lower ($M = 1.0$, $S.E. = .6$) at the *unfriendly*

⁹Agents' eyes were too small to be clearly observed by users since they were approaching them from distance.

level as opposed to the *friendly* level ($M = 1.6$, $S.E. = .6$). The believability of the produced group territorial behavior had been previously validated in Pedica et al. [2010].

In sum, the refined stimuli were adopted to conduct the main study that focused on the measurements described as follows.

5.3. Measures

5.3.1. Subjective: Social Presence. In both trials, we had a subject self-report assessment of the agents' social presence. We simply refer to this measurement as PRESENCE. We took the Temple Presence Inventory (TPI) [Lombard et al. 2011] because it offers standardized measurements for diverse presence dimensions including several types of social presence that were particularly suitable for this study. We selected a relevant subset of dimensions and slightly adapted the questions to our study following the recommendations by Lombard and colleagues. In particular, we measured the following dimensions: *spatial presence* (D.1), *social presence–actor within medium (para-social interaction)* (D.2), *social presence–active interpersonal* (D.3), *engagement* (D.4), *social richness* (D.5), and *social realism* (D.6). Spatial presence (D.1) describes a state of consciousness that gives the impression of being physically present in a mediated world [Lombard and Ditton 1997]. The social presence dimensions (D.2 and D.3) are defined as follows. The actor within medium (para-social interaction) is defined as the apparent face-to-face interaction that can occur between media characters and their audience (i.e., the user in our case) [Short et al. 1976; Ballantine and Martin 2005]. Feelings of para-social interaction are nurtured through carefully constructed mechanisms, such as verbal and nonverbal interaction cues [Labrecque 2014]. According to Lombard and colleagues [2009], users respond to social cues presented by persons they encounter within a medium even though it is illogical or inappropriate to do so, whereas the active interpersonal dimension refers to an anthropomorphism of the medium such that the user experiences the medium/technology itself as a social actor (virtual agents in our case). Engagement (D.4) emphasizes the idea of psychological immersion [Lombard and Ditton 1997]. When users feel psychological immersion they are involved [Palmer 1995] and absorbed [Walker 1998] in the (virtual in our case) experience. Social richness (D.5) measures the extent to which a medium is perceived as sociable, warm, sensitive, personal, or intimate when it is used to interact with other people [Short et al. 1976]. Finally, social realism (D.6) is the extent to which a media portrayal is plausible in that it reflects events that do or could occur in the nonmediated world [Short et al. 1976].

Table II shows the adapted set of questions that we used for each dimension in the JOIN CONVERSATION trial. The questions and anchors used in the REACH DESTINATION trial were the same except for three questions measuring the engagement dimension (D.4). We needed to further adapt them to the different task as follows:

- Q8. How much did you think the group of agents wanted to involve you in their interaction?
- Q9. To what extent did you experience a sensation of reality in the reaction of the agents to you?
- Q10. How relaxing or exciting was completing your task?

5.3.2. Behavioral: Avatar Distance. We also included a behavioral measure named STOP-DISTANCE in the JOIN CONVERSATION trial and PASS-DISTANCE in the REACH DESTINATION trial. The STOP-DISTANCE, as illustrated in Figure 3 on the left, was the distance between the avatar and the group centroid recorded at the moment participants pressed ENTER to complete their task (i.e., join the conversation). The group centroid was a 2D point at the center of an imaginary circumference drawn around the agents (cf. the nucleus as defined by Scheffen [1976]). The PASS-DISTANCE, illustrated on the right of the figure, was recorded as the minimum distance between the

Table II. Questions for the Six Social Presence Dimensions (Abbreviations in the Leftmost Column) Adapted from the Temple Presence Inventory and Used in JOIN CONVERSATION Trial. All Answers Were on a 7-Point Likert Scale. Left and Right Anchors Are Shown in the Rightmost Columns

DIM.	QUESTION	LEFT ANCHOR	RIGHT ANCHOR
D.1	Q1. To what extent did you experience a sense of being there inside the environment with the people you saw/heard?	Not at all	Very much
D.2	Q2. How often did you have the sensation that people you saw/heard could also see/hear you?	Never	Always
	Q3. To what extent did you feel you could interact with the people you saw/heard?	None	Very Much
	Q4. How much did it seem as if you (with your avatar) and the people you saw/heard were together in the same place?	Not at all	Very much
	Q5. How often did it feel as if someone you saw/heard in the environment was looking directly at your avatar?	Never	Always
D.3	Q6. How often did you want to speak to a person you saw/heard in the virtual environment if it was possible?	Never	Always
D.4	Q7. To what extent did you feel mentally immersed in the virtual experience?	Not at all	Very much
	Q8. How involved did you feel in the group interaction?	Not at all	Very much
	Q9. To what extent did you experience a sensation of reality in the interaction with the group?	Not at all	Very much
	Q10. How relaxing or exciting was the interaction with the group of agents?	Very relaxing	Very exciting
D.5	Q11. What is the number that best describes your evaluation of the group of agents?	Unemotional (1)	Emotional (7)
		Unresponsive (1)	Responsive (7)
		Dead (1)	Lively (7)
		Impersonal (1)	Personal (7)
		Insensitive (1)	Sensitive (7)
D.6	Q12. The events I saw/heard would occur in the real world.	Strongly disagree	Strongly agree
	Q13. The events I saw/heard could occur in the real world.	Strongly disagree	Strongly agree
	Q14. The way in which the events I saw/heard occurred is a lot like the way they occur in the real world.	Strongly disagree	Strongly agree

avatar and the group centroid while participants displaced their avatar attempting to go past the group and reach the target destination. In other words, this measured how close to the group they took their avatar while completing the task.

5.3.3. Qualitative: Unstructured Interview. After the completion of the experimental session, we prompted participants open-ended questions about their overall opinions about the experience and suggestions for improvements. We also invited them to reflect on their task and tell us how they would have behaved if it had been assigned in the real world. We aimed at eliciting some qualitative responses about the experience and we also wanted to ensure, for example, in the REACH DESTINATION trial, that participants were not forced to follow a pre-defined path with their avatar but they could choose (i.e., there was enough space) either to pass through the group when

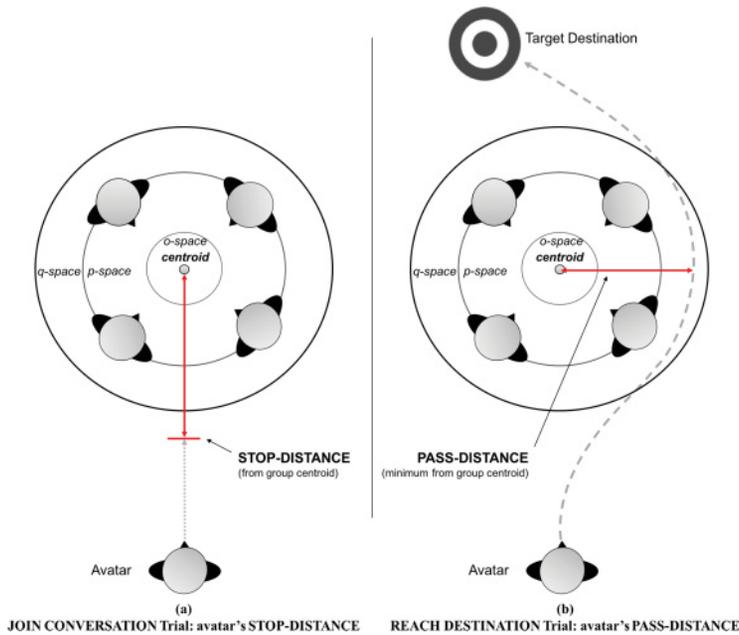


Fig. 3. A schematic illustration of the two behavioral assessments of user's proxemics behavior. The avatar's STOP-DISTANCE in the JOIN CONVERSATION trial (a), and the avatar's (minimum) PASS-DISTANCE in the REACH DESTINATION trial (b).

interpersonal space among agents was higher or on the right side of the group (the agents were placed on the left side of the street).

5.4. Participants and Procedure

5.4.1. Join Conversation Trial. We had 30 participants (12 females and 18 males) recruited at our institute. They were aged 18–60 with 57% in the 21–30 range, and they represented seven nationalities¹⁰ with 57% of them being French; 97% had completed at least undergraduate level education and 50% reported using a video game system at least 1–4 times each month (23% 5–10 times).

5.4.2. Reach Destination Trial. For this trial we had 24 participants (9 females and 15 males). They were aged 18–60 with 58% in the 21–30 range, and they represented seven nationalities with 67% of them being French; 70% had completed at least undergraduate level education and 75% reported using a video game system at least 1–4 times each month (12% more than 20 times).

5.4.3. Procedure. In both trials participants were led to a dedicated room at our laboratory facility, seated in front of a 27" LCD monitor, instructed about the procedure, shown a tutorial for familiarization, and asked to sign an informed consent form. After this introduction, the investigator monitored the session from an adjacent room. The session consisted of (1) completing the task (depending on the trial) for each group of agents ($3 \times 2 = 6$ groups in total) and then filling a computer form that included social presence assessments after each interaction, (2) inserting demographic data in a

¹⁰As part of the demographic information, we asked participants to select the nation that most represented their cultural identity from a list of all countries in the world.

separate form at the end, and (3) finally the investigator debriefed and quickly interviewed them.

The 3×2 levels of our two independent variables led to six conditions shown in a repeated measures within-subjects design. We adopted a Latin-square design to partially **counter balance** the treatment order and **avoid first-order carryover effects** [Bradley 1958].

5.5. Hypotheses

In human-human interaction, expressing a friendly interpersonal attitude increases immediacy between interactants, that is, the perceived physical or psychological closeness between people [Richmond et al. 2008]. In particular, nonverbal expression of friendliness leads to more approachable and welcoming stances [Argyle 1988; Kendon 1990]. In human-agent interaction, previous work confirmed this finding for 1-to-1 greeting encounters [Cafaro et al. 2012], and for this study we hypothesize that this can be extended to 1-to-many interactions where a user approaches a group of agents.

We pose the hypothesis that, for the user, the importance of agents' IN-GROUP and OUT-GROUP attitude depends on the assigned task. Moreover, the behavior exhibited by the agents to obtain the different levels of each variable will influence the social PRESENCE ratings of the group as well as the user's PROXEMICS BEHAVIOR. More specifically, for the JOIN CONVERSATION task, we believe that an expressed OUT-GROUP attitude has greater impact than the IN-GROUP attitude. This means that no matter what attitude the agents express among themselves, even if that behavior is continuously exhibited by the agents and observable by the user, the importance is given to the attitude expressed towards the user (i.e., the approaching avatar in the 3D environment).

On the other hand, in the REACH DESTINATION trial the task is not demanding the user to actively interact with the agents. The conversing group is placed on purpose between the avatar and the target destination, but users can walk around it and might pay more attention to within group behaviors since no direct interaction with the agents is really required and there is more time to observe their behavior (while passing nearby) as opposed to the join conversation trial. Therefore, we believe that in this case the expressed IN-GROUP attitude has more importance and has greater impact on user's subjective judgments and proxemics behavior.

The magnitude of this importance is linked to the specific levels of the two variables and, based on human-human interaction findings mentioned above, we also think that expressing friendly IN-GROUP or OUT-GROUP attitudes will yield to higher ratings of social presence and make the user stop/pass closer to the group of agents, while unfriendly attitudes will produce the effect of "pushing users away" as expressions of noninclusiveness and unwelcoming behavior. To sum up, our hypotheses were the following:

5.5.1. Join Group Trial.

- H1:** The OUT-GROUP attitude of the agents will have a **main effect** (i.e., more important) on subject's PRESENCE evaluation. In particular, participants' PRESENCE evaluation will be higher for *friendly* groups compared to the *unfriendly* ones, regardless of the IN-GROUP attitude of the agents.
- H2:** The OUT-GROUP attitude expressed by the agents will have a **main effect** on subject's proxemics behavior in the 3D world. In particular, participants' STOP-DISTANCE will be closer to *friendly* groups compared to the *unfriendly* ones, regardless of the IN-GROUP attitude of the agents.

Table III. Summary of F -Values, p -Values, and Effect Sizes of RM-ANOVAs Obtained for the Six Presence Dimensions in the JOIN CONVERSATION Trial

Presence Dimension	D.1	D.2	D.3	D.4	D.5	D.6
$F(1,29)$	25.1	32.4	6.3	26.6	37.9	10.6
p -value	.000	.000	.017	.000	.000	.003
Eta Squared	.17	.20	.04	.16	.18	.10
Standardized Mean (S.E.) OUT-GROUP = <i>friendly</i>	.6 (\pm .038)	.58 (\pm .03)	.53 (\pm .034)	.48 (\pm .029)	.5 (\pm .03)	.55 (\pm .035)
Standardized Mean (S.E.) OUT-GROUP = <i>unfriendly</i>	.42 (\pm .033)	.39 (\pm .028)	.45 (\pm .035)	.37 (\pm .024)	.38 (\pm .032)	.43 (\pm .037)

5.5.2. Reach Destination Trial.

- H3:** The IN-GROUP attitude of the agents will have a **main effect** on subject's PRESENCE evaluation. In particular, participants' PRESENCE evaluation will be higher for *friendly* groups compared to the *neutral* ones and, in turn, it will be higher compared to the *unfriendly* ones. This will be **regardless** of the OUT-GROUP attitude expressed by the agents.
- H4:** The IN-GROUP attitude expressed by the agents will have a **main effect** on subject's proxemics behavior in the 3D world. In particular, participants' PASS-DISTANCE will be smaller when passing by *friendly* groups compared to the *neutral* and *unfriendly* ones. This will be regardless of the OUT-GROUP attitude expressed by the agents.

6. RESULTS

6.1. Join Conversation Trial

6.1.1. Subjective: Social Presence. We conducted a two-way repeated-measures MANOVA (i.e., Multivariate repeated measures ANalysis of VAriance) with agents' IN-GROUP and OUT-GROUP attitude as within factors and PRESENCE dimensions as dependent measures. We adopted a multivariate approach since the six presence dimensions in the TPI were moderately inter-correlated with each other ($r \geq 0.5$).

The MANOVA revealed an overall significant main effect of OUT-GROUP attitude on PRESENCE, $Wilks' Lambda = .37$, $F(6, 24) = 6.68$, $p < .001$, $\eta^2 = .626$. Since the sphericity assumption was not violated, we performed a follow-up analysis that looked at univariate effects for each presence dimension with two-way repeated-measures analysis of variance (RM-ANOVAs). These analyses confirmed a significant main effect of OUT-GROUP attitude on all the six presence dimensions measured. All $p \leq .001$ except for *D.3 social presence-active interpersonal* ($p = .017$) and *D.6 social realism* ($p = .003$). Table III shows a summary of all F -values, p -values, and effect sizes. At the bottom of the table and in the chart depicted in Figure 4, the standardized means and errors (in the interval $[0 - 1]$) for the two levels (i.e., *friendly* vs. *unfriendly*) of OUT-GROUP are shown. The main effect of IN-GROUP was not significant. This would suggest that an out group *friendly* attitude overcomes the IN-GROUP attitude among the virtual agents and has greater impact on participants' presence evaluation. In particular, social presence was higher (for all dimensions) when joining a group of agents with *friendly* OUT-GROUP attitude compared to the *unfriendly* groups, thus **H1 is supported**.

No interaction effects were found between the IN-GROUP and OUT-GROUP factors (all $p \geq .21$). We have also looked at possible interaction effects between these two factors and (1) the order of presentation of the six conditions and (2) the participants' nationality. The analysis (1) was done in order to control for possible order effects, though a balanced Latin-square design was deployed, and first-order carryover

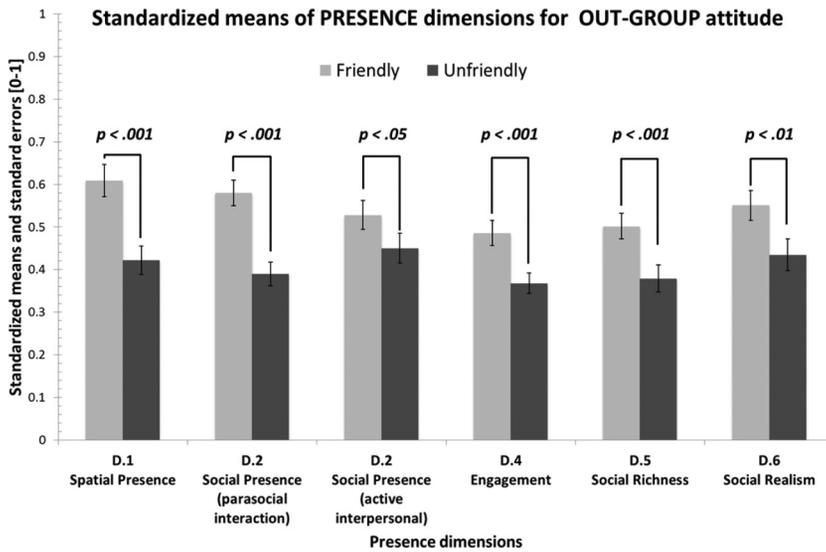


Fig. 4. Standardized means and errors in the interval [0 – 1] for the six PRESENCE dimensions measured in the *friendly* and *unfriendly* OUT-GROUP attitude conditions in the JOIN CONVERSATION trial.

effects should already disappear with this method. We ran a similar RM-ANOVA adding the condition order as between-subjects factor (with six possible levels). The analysis did not reveal any significant interaction effect. The analysis (2) was done for controlling possible cultural differences in the interpretation of virtual agents' nonverbal behavior as pointed out in Endrass et al. [2013]. For this analysis, we ran a similar RM-ANOVA with IN-GROUP and OUT-GROUP as within-subjects factors and participants' nationality as a between-subject factor.¹¹ We did not find significant main effects and interactions to report.

6.1.2. Behavioral: Avatar Stop-Distance. A two-way repeated-measures ANOVA was conducted to compare the effects of IN-GROUP and OUT-GROUP attitudes on participants' STOP-DISTANCE (proxemics behavior) when completing the task. A D'Agostino skewness test [D'Agostino 1970] ran on the original dataset revealed a significant positive skewness in one of our measurements 1.02 ($SE = 2.39$, $p < .05$). In order to obtain a normally distributed dataset and fit the ANOVA model's assumptions, since distances tend to appear with right skewness (i.e., all distances are always > 0 and tend to be log-normally distributed), we applied a \log_{10} – transformation to the original dataset as recommended in Bland and Altman G. [1996].

The RM-ANOVA analysis on log-transformed STOP-DISTANCES revealed a significant interaction effect between IN-GROUP and OUT-GROUP attitudes, $F(2, 58) = 2.54$, $p = .08$, $\eta^2 = .08$. A follow-up analysis revealed that these effects were significant only when OUT-GROUP was *friendly* $F(2, 58) = 5.52$, $p = .006$, $\eta^2 = .16$. A simple main effect analysis was conducted with pairwise comparisons between the different levels of IN-GROUP (when OUT-GROUP was *friendly*) with Bonferroni adjustment for multiple comparisons. The analysis revealed that participants took their avatar closer to groups expressing *friendly* IN-GROUP attitude ($GM = 0.74$, 95% CI [0.6, 0.9]), compared to *neutral* ($GM = 1.02$, 95% CI [0.9, 1.2]) and *unfriendly* groups ($GM =$

¹¹We run a subgroup analysis by created two levels for this factor and assigned the largest group (i.e., *French*) to one level and the remaining participants to the *Other* level.

Table IV. Summary of F -Values, p -Values, and Effect Sizes of RM-ANOVAs Obtained for the Six Presence Dimensions in the REACH DESTINATION Trial

Presence Dimension	D.1	D.2	D.3	D.4	D.5	D.6
$F(1,29)$	21.2	40.1	13.7	37.5	26.1	3.5
p -value	.000	.000	.001	.000	.000	.07
Eta Squared	.16	.30	.11	.35	.28	.06
Standardized Mean (S.E.) OUT-GROUP = <i>friendly</i>	.63 (\pm .04)	.68 (\pm .04)	.53 (\pm .06)	.56 (\pm .03)	.57 (\pm .03)	.58 (\pm .04)
Standardized Mean (S.E.) OUT-GROUP = <i>unfriendly</i>	.52 (\pm .04)	.48 (\pm .04)	.39 (\pm .05)	.40 (\pm .03)	.43 (\pm .03)	.49 (\pm .05)

1.06, 95% CI [0.9, 1.3]).¹² These results are visually summarized in Figure 6 (left) by showing a top view of the group, depicting the concentric areas in which people arrange during a conversation as defined by Schefflen [1976] in the three configurations according to the IN-GROUP attitude levels. This would suggest that our hypothesis **H2** is **partially supported**. There was no main effect of OUT-GROUP attitude as predicted; however, STOP-DISTANCE was influenced by group expressing a *friendly* OUT-GROUP attitude, and participants took their avatar closer to the groups expressing a *friendly* IN-GROUP attitude (depicted in red in the figure).

Similarly to the PRESENCE additional analyses, we looked at possible interaction effects with conditions order of presentation and participants' nationality; however, neither analysis revealed any significant interaction effect.

6.1.3. Qualitative: Unstructured Interview. We conducted unstructured interviews at the end of each session and after having debriefed participants, eliciting their overall experience interacting with the agents across the conditions, suggestions for improvements, and general feedback. We reviewed the interview notes and derived three main themes. First, participants ($n=10$) indicated that a possible suggestion would be to add the “*possibility to talk and interact more with the agents*” once they had accomplished their task (i.e., join the conversation). Second, some participants ($n=4$) pointed out that “*noticing agents' gaze and proxemics behavior was easier than noticing their smiling.*” Finally, no one reported experiencing particular problems due to the interface and the controls of the avatar (i.e., displacement and stopping when close by the group by pressing ENTER).

6.2. Reach Destination Trial

6.2.1. Subjective: Social Presence. We conducted a similar two-way repeated-measures MANOVA for this trial (cf. Section 6.1). The analysis revealed an overall significant main effect of OUT-GROUP attitude on PRESENCE, $Wilks' \Lambda = .308$, $F(6, 18) = 6.75$, $p < .001$, $\eta^2 = .692$.

The follow-up analyses on univariate effects confirmed a significant main effect of OUT-GROUP attitude on all the six presence dimensions measured (all $p \leq .001$) except for *D.6 social realism* ($p = .07$). A summary of the results is shown in Table IV. The standardized means and errors (in the interval [0–1]) for the two levels of OUT-GROUP are shown both at the bottom of the table and in the chart depicted in Figure 5. Contrary to our hypothesis, the main effect of IN-GROUP was not significant, and thus **H3** is **rejected**. The OUT-GROUP attitude had more impact on PRESENCE assessments also in this task. In particular, standardized scores of social presence were higher (for

¹²It should be noted that we show geometric means (GMs) instead of the means, a back transformation of STOP-DISTANCE values was applied to the log-transformed means of the ANOVA test, and this results in geometric and not arithmetic means. The confidence intervals (CI) shown are also back transformed.

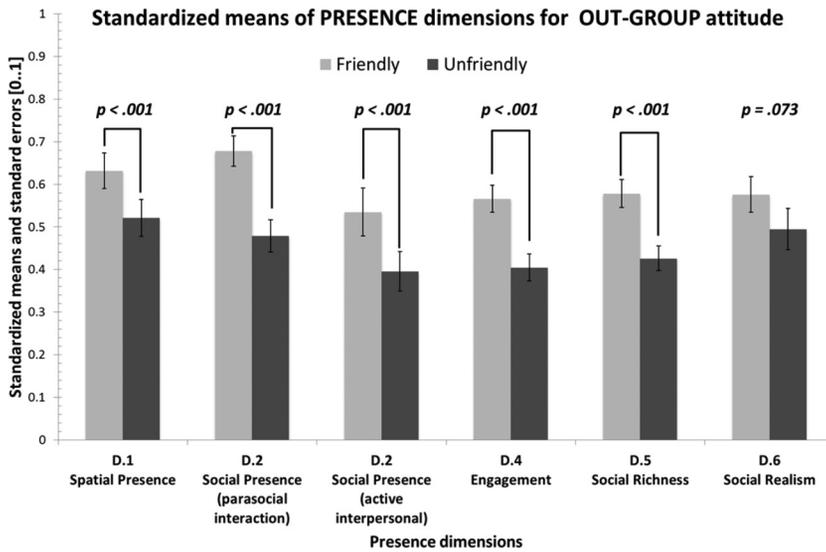


Fig. 5. Standardized means and errors in the interval [0 – 1] for the six PRESENCE dimensions measured in the *friendly* and *unfriendly* OUT-GROUP attitude conditions in the REACH DESTINATION trial.

all dimensions) when participants encountered a group of agents with *friendly* OUT-GROUP attitude on their way to the target destination, compared to the *unfriendly* groups.

No interaction effects were found between the IN-GROUP and OUT-GROUP factors (all $p \geq .17$). We also ran two additional analyses to control for (1) condition order and (2) nationality as between-subjects factors, but also in this case we did not find any significant main or interaction effect.

6.2.2. Behavioral: Avatar Pass-Distance. We performed a two-way repeated-measures ANOVA to compare the effects of IN-GROUP and OUT-GROUP attitudes on participants' PASS-DISTANCE. One of the measurements had significant positive skewness also in this case according to the D'Agostino's test 1.17 ($SE = 2.48$, $p < .05$), thus we \log_{10} -transformed the original dataset.

The RM-ANOVA analysis on log-transformed PASS-DISTANCES revealed a significant interaction effect between IN-GROUP and OUT-GROUP attitudes, $F(2, 46) = 6.8$, $p = .003$, $\eta^2 = .1$. A follow-up analysis revealed that these effects were significant only when OUT-GROUP was *unfriendly* $F(2, 46) = 15.47$, $p < .001$, $\eta^2 = .4$, thus we looked at simple main effects of IN-GROUP (when OUT-GROUP was *unfriendly*). These revealed that participants passed **through** the groups expressing an *unfriendly* IN-GROUP attitude, the geometric (back-transformed) mean was 0.17 (95% CI [0.08, 0.35]), whereas they stayed further away when they were *neutral*, $GM = 0.31$, 95% CI [0.15, 0.67]), and they chose to pass by the group when IN-GROUP attitude was *friendly*, ($GM = 1.15$, 95% CI [0.98, 1.36]).

Figure 6 (right) visualizes these results showing the three symbolic paths of participants' avatars (i.e., dotted lines do not depict the exact paths, and instead they give an idea of the average distance from the groups) for the three levels of IN-GROUP attitude when the OUT-GROUP attitude of the agents was *unfriendly*. It should be noted that the image depicts real mean distances from the group centroid (i.e., not transformed), thus differing from the GM reported above. There was no main effect of

IN-GROUP attitude as predicted; however, PASS-DISTANCE was influenced by both OUT and IN-GROUP attitudes, therefore **H4 is partially supported**.

Similarly to the previous analyses, we looked at possible interaction effects with conditions order of presentation and participants' nationality; however, neither analysis revealed any significant interaction effect.

6.2.3. Qualitative: Unstructured Interview. The unstructured interviews in this trial revealed several interesting themes. First, it was reported (n=2) that “*when the agents were moving all together towards my avatar the group seemed a bit aggressive towards me*” (i.e., this happened when the OUT-GROUP attitude was *friendly*, thus the group was opening upon detecting the avatar approaching close by). Furthermore, some participants (n=5) “*attempted to join the group but it was not possible*” even if the task assigned was to reach the target destination, and two participants reported to have “*looked back at the group once reached the target to check if they were looking at me.*” Second, we asked all participants if they passed with their avatar through the group of agents. For those who did it, we further asked if they felt there was no space to navigate with their avatar by the group's right side, so they were forced to pass through the center of the group; however, all of them responded that it was their choice to go either on the right side or centrally but space was not an issue. A couple of participants reported having some difficulties using the controls though, because they were not familiar with gaming environments. Finally, certain participants (n=6) found out that “*the last three questions in the questionnaire were very similar by adopting a wording usage too subtle to differentiate*”; this referred to D.6's questions (i.e., social realism) and the use of the modal verbs “*would*” and “*could*” (cf. questions Q12–Q14 in Table II). Questions Q12–Q14 wording's might have generated confusion in the participants, thus explaining the lack of significance difference in the ratings relative to D.6. In sum, these qualitative findings provided useful insights on the overall participants' experience that we will discuss in the next section together with all the results obtained.

7. DISCUSSION

The results when users had the task of joining a conversation supported **H1**. We found that users attributed higher social presence to group of agents expressing a *friendly* OUT-GROUP attitude towards them compared to *unfriendly* one, no matter what IN-GROUP attitude the agents had among themselves in the simulated conversation. In particular, all presence dimensions measured were rated significantly higher in the *friendly* OUT-GROUP conditions compared to the *unfriendly* ones. The effect size (shown in Table III) was particularly pronounced for the dimensions **D.1**, **D.2**, **D.4**, and **D.5** characterizing (cf. Section 5.3.1 at page 16), respectively, the **spatial presence** (D.1), the feeling of **para-social interaction** (D.2), the level of **engagement** (D.4) with the agents, and the **social richness** (D.5) expressed by the conversing groups.

When looking into greater detail to these dimensions, we observe that they constitute perceptions such as the feeling of being in the same place (D.1), the feeling of having a natural interaction with the group in response to the observed nonverbal social cues exhibited towards the avatar (D.2), and an increased level of psychological involvement (D.4) and the judgments related to the extent the groups were perceived as sociable and warm (D.5). These effects were due to the expressed level of OUT-GROUP friendliness, which in turn (as validated by our manipulation check, cf. Section 5.2.1), was obtained through the nonverbal behavior exhibited by the agents. It is possible that instances of specific behaviors, such as gaze, had greater importance for certain dimensions. When decomposing D.2 (para-social interaction), for example, we notice that some of the questions in the inventory rely on the perceived amount of gaze behavior received by the agents. However, the behavior exhibited by the agents had also impact on other

dimensions assessed differently (e.g., D.4 and D.5) that capture different aspects of the interaction (i.e., social richness and engagement) and are not explicitly linked to specific behaviors. We believe that these effects were obtained by exhibiting coordinated nonverbal behaviors (including proxemics and smiling) viewed as a whole. So far, we know very little about the relative importance of specific nonverbal cues exhibited in isolation from the others for expressing group attitudes. This represents an interesting aspect to investigate in further similar studies (see Section 9).

Another possible interpretation of these outcomes is offered by looking deeper in the nonverbal cues exhibited by our groups for expressing their out-group attitudes. More specifically, eye contact, smiling, and interpersonal distance are some of the main nonverbal cues contributing to **immediacy**, defined as the degree of perceived physical or psychological closeness between people [Richmond et al. 2008]. There are several positive outcomes of immediacy in interpersonal interaction. First, immediacy increases liking and affiliation (i.e., friendliness) [Mehrabian 1969b]. This directly relates to our manipulations since our primary goal was to manipulate the level of out-group friendliness of our agents. Second, the concept of immediacy is grounded in the approach-avoidance theory. According to Mehrabian [1981], someone exhibiting immediacy is perceived as having a more approachable communication style and people approach what they like and avoid what they do not like. Third, immediacy behaviors are considered to be highly effective contributors for perceived enthusiasm, expressiveness, and responsiveness [Richmond et al. 2008]. Moreover, eye contact and proxemics are great contributors to engagement in interaction; closer people exchanging more eye contact are perceived as more engaged in interaction compared to people exhibiting nonimmediate nonverbal behavior [Tatar 1997]. The notion of social presence adopted in this article is a combination of these factors. This suggests that greater intimacy between a group and a newcomer (i.e., the user in our study) has a positive effect on the newcomer's affective filters and ultimately impacting the different dimensions of presence corresponding to the positive outcomes of immediacy. On the other hand, the expression of unfriendly attitude through subtle changes, such as lower amount of gaze at the user and noninclusive postural cues (i.e., indirect body orientation towards the avatar) led to lower immediacy between the user and the group. The users completed their task (i.e., join the conversation) because it was required; however, their subjective evaluation of the group through the presence questionnaire revealed an overall negative judgment (e.g., unsociable and unemotional) due to the sense of being ignored and excluded while still being noticed by the agents.

In sum, coloring the exhibited behavior of the agents for managing their group formation with attitude expression was still perceived as plausible and had significant impacts on user's ratings of social presence dimensions. The *unfriendly* groups were perceived as noninclusive and unsociable, and users reflected their sense of being socially excluded into their lower ratings of social presence. Moreover, we believe that users perceived the observed agents behavior and reaction as believable, since they reported being willing to interact more with the agents if that was possible (cf. qualitative findings in Section 6.1).

The results of user's proxemics behavior (i.e., STOP-DISTANCE of their avatar) partially supported **H2**. The OUT-GROUP attitude played a role also in this case but in interaction with the IN-GROUP attitude of the agents. As depicted in Figure 6 (left), the STOP-DISTANCE¹³ significantly differed between levels of IN-GROUP attitude when the OUT-GROUP attitude was *friendly* (i.e., group members orienting their body

¹³The values in the image are the real mean distances (i.e., not the geometric means) measured from the group centroid and represented in 3D units, that is, 1 unit in the virtual environment is approximately 1 real-world meter.

towards users) sufficed to significantly impact the presence judgments. The impact of OUT-GROUP attitude on presence judgments can be explained by the participants' curiosity in the synthetic characters, also supported by our qualitative findings, when participants reported attempting to join the group even if not required by their task, and checking if the agents were looking at their avatar once reached the target destination. Therefore, people have high expectations for agents' behavior. Even if in small doses, the agents' reactions towards their avatar had impact on the social presence attributed to the group of characters.

The user's proxemics behavior (i.e., PASS-DISTANCE) was influenced by both OUT-GROUP and IN-GROUP attitudes, thus **partially** supporting **H4**. Contrary to our expectations, however, the unfriendly OUT-GROUP distance had an impact on this behavioral measure and interacted with the IN-GROUP attitudes. As depicted in Figure 6 (right), participants passed through the group when the agents were unfriendly among themselves and chose to go by the side in the neutral and friendly cases. Participants reported perceiving as aggressive the groups attempting to open and move towards their avatar (i.e., *friendly* OUT-GROUP), thus possibly explaining why they had less of a problem passing through the *unfriendly* OUT-GROUP. The different distances among the agents (as function of the IN-GROUP attitude) might be the additional explanation for this outcome. When *unfriendly*, the agents stood farther away from each other, thus making space for the avatars to pass through, although participants felt they had space on both sides (cf. qualitative results in Section 6.2).

Earlier work highlighted the importance of considering social norms and awareness of territories, such as the space between those having a conversation, for obtaining believable small conversing group simulations [Oliva and Vilhjálmsson 2014]. Previous findings also revealed that humans process and simplify interactions by considering a group of many as a whole. They choose to pass through large sparsed groups and to avoid (going around) small densed groups [Bruneau et al. 2015]. This behavioral choice is in line with the principle of the least effort, stating that people naturally choose the path of least resistance or effort [Kluckhohn 1950]. Thus, in our study participants (a) adhered to the group's social norms while finding their path to the destination and (b) chose the simplest path for achieving their task.

We should also remark that OUT-GROUP attitude is observable earlier during the avatar's approach to the group, and since it had a major impact on presence dimensions' ratings, the lower rating of an unfriendly OUT-GROUP might have lowered the pressure to observe social norms and resulted in users passing directly through groups. These results are also in line with the experiments described in Ennis and O'Sullivan [2012], where participants were more sensitive to agents' distance and orientation changes in a conversing group, compared to anomalous gesturing behavior in terms of perceived believability.

Overall, these two trials demonstrated that the agents OUT-GROUP attitude impacts users' social presence ratings more than the IN-GROUP attitude expressed among group members. The behavior exhibited by the agents to express both in and out group attitudes subtly differed, but users were more sensible to those behaviors directed to them (i.e., out-group), even when their task was to reach a given destination and did not require any direct interaction with the group. The believability of the simulated conversations held, even when the autonomous behavior generated by our system, for reproducing a plausible group formation and territorial behavior, was modulated for expressing social attitudes. It should be noted that the assigned tasks, although very common in simulation and gaming environments, mainly focused on avatar locomotion in the 3D environment. Therefore the interpersonal distance among the agents and between the whole group and their avatar played an important role, in particular for our behavioral assessments (i.e., STOP-DISTANCE and PASS-DISTANCE).

There was not enough cultural variation in both samples of participants. We obtained rather homogeneous samples (i.e., the majority of participants were French), thus the effects that we found can be attributed to individual differences rather than cultural variations. However, further investigation can be conducted for establishing whether the relationship and combined effects of IN-GROUP and OUT-GROUP attitudes are culture independent even if the social parameters for expressing them vary across cultures.

In the physical world, nonverbal reactions, such as a smile, gaze behavior, and interpersonal distance are powerful signals for indicating the desire to interact [Argyle 1988]. Smiling is a common feature in greeting rituals [Kendon 1990] and a sign of positive affect [Richmond et al. 2008]. Interpersonal distance and body orientation are important body cues for expressing social attitudes, for example, greater distance indicates the desire for greater formality, while direct body orientation, openness, and closer distances increase immediacy (i.e., friendliness) [Argyle 1988]. From this study, we observed that these norms migrated into the 3D world, and, more specifically, our system was capable of autonomously generating the agents' nonverbal behavior supporting an expressive and believable simulated group conversation.

7.1. Guidelines

Our quantitative and qualitative findings suggest that users accepted the exhibited group behavior and perceived the synthetic characters as believable. However, if the ultimate goal is to use the presented technology (or similar systems) for populating virtual worlds, making these worlds more believable, and donating social intelligence and expressive traits to the virtual agents, our results have some implications for the design and implementation of such systems:

- An interactive simulated environment populated by virtual humans can rely on autonomous generation of agents' multimodal behavior to increase the believability. In a scenario where the user's experience is built on the possibility of interacting with these agents (i.e., avatar based interaction), particular attention should be given to the outward behavior of the agents, while simpler techniques can be used to simulate the behavior among the agents (e.g., not requiring the usage of expensive motion capture based data to animate their in-group behavior).
- In scenarios where the agents are part of the background story and their attitude is important for conveying certain traits (such as an unfriendly stance during a group dispute), nonverbal behavior similar to that generated by our system can be exhibited by the agents, while keeping in mind that portrayal of unfriendliness may lead to a reduced sense of perceived social presence and believability.

8. CONCLUSIONS

In this article, we described an integrated intelligent system for generating believable nonverbal behavior exhibited by virtual agents in simulated small group conversations. The produced nonverbal behavior supported group formation management and the expression of friendly/unfriendly attitudes both among the agents in the group (in-group) and towards an approaching user in an avatar-based interaction (out-group).

We conducted a user study investigating the effects of these attitudes on users' felt sense of social presence and proxemics behavior (with their avatar) in a 3D virtual city environment. In particular, we designed a two-trial within-subjects experiment based on typical scenarios of 3D gaming environments. In a first trial the task was approaching and joining the group of agents, and in the second it was to reach a target destination behind the group. We found a main effect of out-group attitude on presence

evaluations in both trials: Friendly groups were perceived as more socially rich and believable, regardless of the in-group attitude as well as when the task did not require the users to directly interact with the agents (i.e., 2nd trial). The users' proxemics behavior depended on both out-group and in-group attitudes expressed by the agents.

9. LIMITATIONS AND FUTURE WORK

Some limitations of our system should be considered. The expression of group interpersonal attitude was obtained with a variety of nonverbal human-inspired behavior. It would be interesting to conduct a breakdown study focusing on separate cues to understand the most prominent one(s) for expressing friendliness, for example. Furthermore, the agents' exhibited behavior could be enriched for handling important conversational dynamics such as turn-taking strategies (e.g., gaze behavior) and back-channel mechanisms (e.g., head nods). Then, the produced behavior could be altered by our system for expressing specific attitudes. For example, a friendly listener agent could smile and nod more in reaction to the group's speaker compared to an unfriendly one. Moreover, the status dimension of our attitude model should be considered and studied in future work. Dominance, for example, can also be a determinant for social presence judgments in a simulated conversation.

The experimental design has some limitations, too. We had a fixed number of agents in the conversing group and all of them were expressing the same level of attitude in a given condition. While this is plausible in real life, it is also possible that certain members might express friendliness, whereas others might be more distant and unfriendly in a conversation, for instance. We believe that dealing with different levels of attitudes or manipulating the number of participants in the conversation (as a new independent variable) are challenging issues. On one hand, they are interesting aspects to study, but, on the other hand, they make the experimental design more complex. Moreover, we created the conditions using specific characters' appearances and designed the participants' tasks with common scenarios of gaming environments in mind. The character's visual appearance and the assigned tasks are important aspects that should be considered when building on the current study's results.

We plan to continue this line of investigation on the expression of social attitudes, possibly modeling both dimensions of attitude and different levels expressed at the same time among the agents and towards the users. We also envision to enrich the set of behaviors displayed by our agents, including, for example, the support for different turn taking strategies during the conversation reflecting the interpersonal attitude of the participants.

Finally, we aim to deploy and test our system in a different presentation format. Given the increased ubiquity of modern immersive environments technology and more accessible costs (e.g., Oculus Rift device), it will be of primary importance to harness these new technologies and have a better understanding of the impact that such immersive experiences might have on users. We intend to adopt these technologies and consider immersive reality for designing and evaluating new compelling scenarios in future developments of our system. Virtual worlds, for example, have been already used to study social exclusion (i.e., ostracism) [Kassner et al. 2012], but it remains a question whether an unfriendly group of virtual humans is the equivalent of an ostracizing group for a newcomer.

ACKNOWLEDGMENTS

We are grateful to Claudio Pedica for granting us the use of the Impulsion Social AI Engine and providing technical support during the integration of the engine in our system.

REFERENCES

- Michael Argyle. 1988. *Bodily Communication* (2nd ed.). Methuen, New York. 363 pages.
- Michael Argyle, Florisse Alkema, and Robin Gilmour. 1971. The communication of friendly and hostile attitudes by verbal and non-verbal signals. *Eur. J. Soc. Psychol.* 1, 3 (1971), 385–402. DOI: <http://dx.doi.org/10.1002/ejsp.2420010307>
- Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2003. Interpersonal distance in immersive virtual environments. *Person. Soc. Psychol. Bull.* 29, 7 (2003), 819–833.
- Paul Ballantine and Brett Martin. 2005. Forming parasocial relationships in online communities. *Adv. Consumer Res.* 32 (2005), 197.
- Daniel Ballin, Marco Gillies, and Barry Crabtree. 2004. A framework for interpersonal attitude and non-verbal communication in improvisational visual media production. In *First European Conference on Visual Media Production IEEE*. 203–210.
- Aryel Beck, Brett Stevens, Kim A. Bard, and Lola Cañamero. 2012. Emotional body language displayed by artificial agents. *ACM Trans. Interact. Intell. Syst.* 2, 1, Article 2 (March 2012), 29 pages. DOI: <http://dx.doi.org/10.1145/2133366.2133368>
- S. Bernard, J. Therien, C. Malone, S. Beeson, A. Gubman, and R. Pardo 2008. Taming the mob: Creating believable crowds in assassin's creed. In *Presented at Game Developers Conference (San Francisco, CA, Feb 18-22)*.
- Elisabetta Bevacqua, Sathish Pammi, Sylwia Julia Hyniewska, Marc Schröder, and Catherine Pelachaud. 2010. Multimodal backchannels for embodied conversational agents. In *Proceedings of the 10th International Conference on Intelligent Virtual Agents*. Springer, Berlin, 194–200. DOI: http://dx.doi.org/10.1007/978-3-642-15892-6_21
- Martin J. Bland and G. Douglas Altman 1996. Statistics notes: Transforming data. *Br. Med. J.* 312, 7033 (3 1996), 770.
- James V. Bradley. 1958. Complete counterbalancing of immediate sequential effects in a latin square design. *J. Am. Stat. Assoc.* 53, 282 (1958), 525–528.
- Julien Bruneau, Anne-Helene Olivier, and Julien Pettre. 2015. Going through, going around: A study on individual avoidance of groups. *IEEE Trans. Vis. Comput. Graph.* 21, 4 (2015), 520–528.
- Judee Burgoon and Beth Le Poire. 1999. Nonverbal cues and interpersonal judgments: Participant and observer perceptions of intimacy, dominance, composure, and formality. *Commun. Monogr.* 66, 2 (1999), 105–124.
- Judee K. Burgoon, David B. Buller, Jerold L. Hale, and Mark A. de Turck. 1984. Relational messages associated with nonverbal behaviors. *Hum. Commun. Res.* 10, 3 (1984), 351–378.
- Judee K. Burgoon and Jerold L. Hale. 1984. The fundamental topoi of relational communication. *Commun. Monogr.* 51, 3 (1984), 193–214.
- Angelo Cafaro, Hannes Högni Vilhjálmsson, Timothy W. Bickmore, Dirk Heylen, and Daniel Schulman. 2013. First impressions in user-agent encounters: The impact of an agent's nonverbal behavior on users' relational decisions. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, 1201–1202.
- Angelo Cafaro, Hannes Vilhjálmsson, Timothy Bickmore, Dirk Heylen, Kamilla Jóhannsdóttir, and Gunnar Valgarðsson. 2012. First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In *Proceedings of the 12th International Conference on Intelligent Virtual Agents*. Springer-Verlag, Berlin, 67–80.
- Mathieu Chollet, Magalie Ochs, and Catherine Pelachaud. 2014. From non-verbal signals sequence mining to Bayesian networks for interpersonal attitudes expression. In *Proceedings of the 14th International Conference on Intelligent Virtual Agents*, Timothy Bickmore, Stacy Marsella, and Candace Sidner (Eds.). Lecture Notes in Computer Science, Vol. 8637. Springer, Berlin, 120–133.
- Marco Cristani, Giulia Paggetti, Alessandro Vinciarelli, Loris Bazzani, Gloria Menegaz, and Vittorio Murino. 2011. Towards computational proxemics: Inferring social relations from interpersonal distances. In *2011 IEEE 3rd International Conference on Social Computing*. 290–297.
- Ralph B. D'Agostino. 1970. Transformation to normality of the null distribution of g_1 . *Biometrika* 57, 3 (1970), 679–681.
- Starkey Duncan. 1972. Some signals and rules for taking speaking turns in conversations. *J. Person. Soc. Psychol.* 23, 2 (1972), 283.
- Birgit Endrass, Elisabeth André, Matthias Rehm, and Yukiko Nakano. 2013. Investigating culture-related aspects of behavior for virtual characters. *Auton. Agents Multi-Agent Syst.* 27, 2 (2013), 277–304. DOI: <http://dx.doi.org/10.1007/s10458-012-9218-5>

- Cathy Ennis, Rachel McDonnell, and Carol O'Sullivan. 2010. Seeing is believing: Body motion dominates in multisensory conversations. *ACM Trans. Graph.* 29, 4, Article 91 (July 2010), 9 pages. DOI : <http://dx.doi.org/10.1145/1778765.1778828>
- Cathy Ennis and Carol O'Sullivan. 2012. Perceptually plausible formations for virtual conversers. *Comput. Anim. Virtual Worlds* 23, 3–4 (May 2012), 321–329. DOI : <http://dx.doi.org/10.1002/cav.1453>
- Doron Friedman, Anthony Steed, and Mel Slater. 2007. Spatial social behavior in second life. In *Proceedings of the 7th International Conference on Intelligent Virtual Agents*, Catherine Pelachaud, Jean-Claude Martin, Elisabeth André, Gérard Chollet, Kostas Karpouzis, and Danielle Pelé (Eds.). Lecture Notes in Computer Science, Vol. 4722. Springer, Berlin, 252–263.
- Edward T. Hall. 1966. *The Hidden Dimension*. Doubleday.
- Kotaro Hayashi, Masahiro Shiomi, Takayuki Kanda, Norihiro Hagita, and ATR Intelligent Robotics. 2012. Friendly patrolling: A model of natural encounters. In *Proceedings of Robotics: Science and Systems*. 121.
- Dirk Helbing and Péter Molnár. 1995. Social force model for pedestrian dynamics. *Phys. Rev. E* 51 (1995), 4282–4286. Issue 5.
- Dirk Heylen, Stefan Kopp, Stacy C. Marsella, Catherine Pelachaud, and Hannes Vilhjálmsón. 2008. The next step towards a function markup language. In *Proceedings of the 8th International Conference on Intelligent Virtual Agents (IVA'08)*. Springer-Verlag, Berlin, 270–280. DOI : http://dx.doi.org/10.1007/978-3-540-85483-8_28
- Dušan Jan and David R. Traum. 2007. Dynamic movement and positioning of embodied agents in multiparty conversations. In *Proceedings of the Workshop on Embodied Language Processing (EmbodiedNLP'07)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 59–66.
- Kristiina Jokinen, Hirohisa Furukawa, Masafumi Nishida, and Seiichi Yamamoto. 2013. Gaze and turn-taking behavior in casual conversational interactions. *ACM Trans. Interact. Intell. Syst.* 3, 2, Article 12 (Aug. 2013), 30 pages. DOI : <http://dx.doi.org/10.1145/2499474.2499481>
- Matthew P. Kassner, Eric D. Wesselmann, Alvin Ty Law, and Kipling D. Williams. 2012. Virtually ostracized: Studying ostracism in immersive virtual environments. *J. Cyberpsychol. Behav. Soc. Network.* 15, 8 (2012), 399–403.
- Adam Kendon. 1967. Some functions of gaze-direction in social interaction. *Acta Psychol.* 26, 1 (1967), 22–63. DOI : [http://dx.doi.org/10.1016/0001-6918\(67\)90005-4](http://dx.doi.org/10.1016/0001-6918(67)90005-4)
- Adam Kendon. 1990. *Conducting Interaction: Patterns of Behavior in Focused Encounters (Studies in Interactional Sociolinguistics)*. Cambridge University Press, Cambridge.
- Yunkyung Kim and Bilge Mutlu. 2014. How social distance shapes human–robot interaction. *International J. Hum.-Comput. Stud.* 72, 12 (2014), 783–795. DOI : <http://dx.doi.org/10.1016/j.ijhcs.2014.05.005>
- Nathan Kirchner, Alen Alempijevic, and Gamini Dissanayake. 2011. Nonverbal robot-group interaction using an imitated gaze cue. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI'11)*. ACM, New York, NY, 497–504. DOI : <http://dx.doi.org/10.1145/1957656.1957824>
- Felix Kistler, Birgit Endrass, Ionut Damian, ChiTai Dang, and Elisabeth André. 2012. Natural interaction with culturally adaptive virtual characters. *J. Multimodal User Interf.* 6, 1–2 (2012), 39–47. DOI : <http://dx.doi.org/10.1007/s12193-011-0087-z>
- Clyde Kluckhohn. 1950. Human behavior and the principle of least effort. George Kingsley Zipf. *Am. Anthropol.* 52, 2 (1950), 268–270.
- Stefan Kopp, Brigitte Krenn, Stacy Marsella, Andrew N. Marshall, Catherine Pelachaud, Hannes Pirker, Kristinn R. Thórisson, and Hannes H. Vilhjálmsón. 2006. Towards a common framework for multimodal generation: The behavior markup language. In *Proceedings of the 6th International Conference on Intelligent Virtual Agents (IVA'06)*. Springer-Verlag, Berlin, 205–217.
- Lauren I. Labrecque. 2014. Fostering consumer–Brand relationships in social media environments: The role of parasocial interaction. *J. Interact. Market.* 28, 2 (2014), 134–148. DOI : <http://dx.doi.org/10.1016/j.intmar.2013.12.003>
- Petri Lankoski. 2010. *Character-Driven Game Design: A Design Approach and Its Foundations in Character Engagement*. Vol. 101. Taik Books.
- Jina Lee and Stacy Marsella. 2011. Modeling side participants and bystanders: The importance of being a laugh track. In *Proceedings of the 11th International Conference on Intelligent Virtual Agents*, Hannes Högni Vilhjálmsón, Stefan Kopp, Stacy Marsella, and Kristinn R. Thórisson (Eds.). Lecture Notes in Computer Science, Vol. 6895. Springer, Berlin, 240–247. DOI : http://dx.doi.org/10.1007/978-3-642-23974-8_26
- Joan Llobera, Bernhard Spanlang, Giulio Ruffini, and Mel Slater. 2010. Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Trans. Appl. Percept.* 8, 1, Article 3 (2010), 12 pages.

- Matthew Lombard and Theresa Ditton. 1997. At the heart of it all: The concept of presence. *J. Comput.-Mediat. Commun.* 3, 2 (1997), 0–0. DOI: <http://dx.doi.org/10.1111/j.1083-6101.1997.tb00072.x>
- Matthew Lombard, Lisa Weinstein, and Theresa Ditton. 2011. Measuring telepresence: The validity of the temple presence inventory (TPI) in a gaming context. In *2011 Annual conference of the International Society for Presence Research (ISPR) (ISPR 2011)*. 9.
- T. Lombard, M. Ditton, and L. Weinstein. 2009. Measuring telepresence: The temple presence inventory. In *Twelfth International Workshop on Presence*.
- David McNeill. 1992. *Hand and Mind*. University of Chicago Press, Chicago IL.
- Ross Mead, Amin Atrash, and Maja J. Matarić. 2013. Automated proxemic feature extraction and behavior recognition: Applications in human-robot interaction. *Int. J. Soc. Robot.* 5, 3 (2013), 367–378. DOI: <http://dx.doi.org/10.1007/s12369-013-0189-8>
- Albert Mehrabian. 1969a. Significance of posture and position in the communication of attitude and status relationships. *Psychol. Bull.* 71, 5 (1969), 359–372.
- Albert Mehrabian. 1969b. Some referents and measures of nonverbal behavior. *Behav. Res. Methods Instrum.* 1, 6 (1969), 203–207. DOI: <http://dx.doi.org/10.3758/BF03208096>
- Albert Mehrabian. 1981. *Silent Messages: Implicit Communication of Emotions and Attitudes*. Wadsworth Publishing Company.
- Bilge Mutlu, Takayuki Kanda, Jodi Forlizzi, Jessica Hodgins, and Hiroshi Ishiguro. 2012. Conversational gaze mechanisms for humanlike robots. *ACM Trans. Interact. Intell. Syst.* 1, 2, Article 12 (Jan. 2012), 33 pages. DOI: <http://dx.doi.org/10.1145/2070719.2070725>
- Nasser Nassiri, Norman Powell, and David Moore. 2010. Human interactions and personal space in collaborative virtual environments. *Virtual Reality* 14, 4 (2010), 229–240.
- Carmine Oliva and Hannes Högni Vilhjálmsón. 2014. Prediction in social path following. In *Proceedings of the Seventh International Conference on Motion in Games (MIG'14)*. ACM, New York, NY, 103–108. DOI: <http://dx.doi.org/10.1145/2668064.2668103>
- Mark T. Palmer. 1995. Communication in the age of virtual reality. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, Chapter Interpersonal Communication and Virtual Reality: Mediating Interpersonal Relationships, 277–299. <http://dl.acm.org/citation.cfm?id=207922.207932>
- Florian Pecune, Angelo Cafaro, Mathieu Chollet, Pierre Philippe, and Catherine Pelachaud. 2014. Suggestions for extending SAIBA with the VIB platform. In *Workshop on Architectures and Standards for IVAs, Held at the 14th International Conference on Intelligent Virtual Agents (IVA 2014)*. Bielefeld eCollections, 16–20.
- Claudio Pedica and Hannes Högni Vilhjálmsón. 2010. Spontaneous avatar behavior for human territoriality. *Appl. Artif. Intell.* 24, 6 (2010), 575–593.
- Claudio Pedica, Hannes Högni Vilhjálmsón, and Marta Lárusdóttir. 2010. Avatars in conversation: The importance of simulating territorial behavior. In *Proceedings of the 10th International Conference on Intelligent Virtual Agents*, Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud, and Alla Safonova (Eds.). Lecture Notes in Computer Science, Vol. 6356. Springer, Berlin, 336–342.
- Claudio Pedica and Hannes Högni Vilhjálmsón. 2012. Lifelike interactive characters with behavior trees for social territorial intelligence. In *ACM SIGGRAPH 2012 Posters (SIGGRAPH'12)*. ACM, New York, NY, Article 32, 1 pages. DOI: <http://dx.doi.org/10.1145/2342896.2342938>
- Rui Prada and Ana Paiva. 2005. Believable groups of synthetic characters. In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'05)*. ACM, New York, NY, 37–43.
- Brian Ravenet, Angelo Cafaro, Beatrice Biancardi, Magalie Ochs, and Catherine Pelachaud. 2015. Conversational behavior reflecting interpersonal attitudes in small group interactions. In *Proceedings of the 15th International Conference on Intelligent Virtual Agents*, Willem-Paul Brinkman, Joost Broekens, and Dirk Heylen (Eds.). Lecture Notes in Computer Science, Vol. 9238. Springer, Berlin, 375–388. DOI: http://dx.doi.org/10.1007/978-3-319-21996-7_41
- Brian Ravenet, Angelo Cafaro, Magalie Ochs, and Catherine Pelachaud. 2014. Interpersonal attitude of a speaking agent in simulated group conversations. In *Proceedings of the 14th International Conference on Intelligent Virtual Agents*, Timothy Bickmore, Stacy Marsella, and Candace Sidner (Eds.). Lecture Notes in Computer Science, Vol. 8637. Springer, Berlin, 345–349. DOI: http://dx.doi.org/10.1007/978-3-319-09767-1_45
- Brian Ravenet, Magalie Ochs, and Catherine Pelachaud. 2013. From a user-created corpus of virtual agent's non-verbal behavior to a computational model of interpersonal attitudes. In *Proceedings of the 13th International Conference on Intelligent Virtual Agents*, Ruth Aylett, Brigitte Krenn, Catherine Pelachaud, and Hiroshi Shimodaira (Eds.). Lecture Notes in Computer Science, Vol. 8108. Springer, Berlin, 263–274.

- Matthias Rehm and Birgit Endrass. 2009. Rapid prototyping of social group dynamics in multiagent systems. *AI Soc.* 24, 1 (2009), 13–23.
- Craig Reynolds. 1999. Steering behaviors for autonomous characters. In *Proceedings of the Game Developers Conference*. Miller Freeman Game Groups, San Francisco, CA, 763–782.
- Virginia Richmond, James McCroskey, and Mark Hickson. 2008. *Nonverbal Communication in Interpersonal Relations* (6th ed. ed.). Allyn & Bacon, London.
- A. E. Schefflen. 1976. *Human Territories: How we Behave in Space and Time*. Prentice-Hall, New York, NY.
- W. C. Schutz. 1958. *FIRO: A Three-Dimensional Theory of Interpersonal Behaviour*. Holt, Rinehart and Winston, New York, NY. 267 pages.
- John Short, Ederyn Williams, and Bruce Christie. 1976. *The Social Psychology of Telecommunications*. John Wiley & Sons Ltd., New York, NY.
- Mel Slater, Amela Sadagic, Martin Usoh, and Ralph Schroeder. 2000. Small-group behavior in a virtual and real environment: A comparative study. *Presence: Teleoperat. Virtual Environ.* 9, 1 (Feb. 2000), 37–51. DOI : <http://dx.doi.org/10.1162/105474600566600>
- Lynn Smith-Lovin and Charles Brody. 2009. *Interruptions in Group Discussions: The Effects of Gender and Group Composition*. SAGE Publications Ltd., London, 424–435. DOI : <http://dx.doi.org/10.4135/9781446262269>
- Olivier Szymanczyk, Patrick Dickinson, and Tom Duckett. 2011. From individual characters to large crowds: Augmenting the believability of open-world games through exploring social emotion in pedestrian groups. In *Proceedings of the 2011 DiGRA International Conference: Think Design Play*. DiGRA/Utrecht School of the Arts.
- D. G. Tatar. 1997. *Social and Personal Effects of a Preoccupied Listener*. Stanford University.
- J. Vroon, M. Joosse, M. Lohse, J. Kolkmeier, Jaebok Kim, K. Truong, G. Englebienne, D. Heylen, and V. Evers. 2015. Dynamics of social positioning patterns in group-robot interactions. In *24th IEEE International Symposium on Robot and Human Interactive Communication*. 394–399. DOI : <http://dx.doi.org/10.1109/ROMAN.2015.7333633>
- G. J. Walker. 1998. *Our Sweetest Hours: Recreation and the Mental State of Absorption*. Taylor & Francis, London.
- Nick Yee, Jeremy N. Bailenson, Mark Urbanek, Francis Chang, and Dan Merget. 2007. The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *Cyberpsychol. Behav.* 10, 1 (2007), 115–121.
- Victor H. Yngve. 1970. On getting a word in edgewise. In *Chicago Linguistics Society, 6th Meeting*. 567–578.

Received June 2015; revised January 2016; accepted April 2016