Don't Walk Between Us: Adherence to Social Conventions When Joining a Small Conversational Group of Agents

Alessandro Iop
KTH Royal Institute of Technology
Stockholm, Sweden
aiop@kth.se

Sahba Zojaji
KTH Royal Institute of Technology
Stockholm, Sweden
zojaji@kth.se

Christopher Peters
KTH Royal Institute of Technology
Stockholm, Sweden
chpeters@kth.se

ABSTRACT

When modeling life-like Embodied Conversational Agents (ECAs), conveying politeness through verbal and nonverbal behaviors with persuasive intents is a significant challenge, as it underlies the conventional set of behavioral rules that govern human communication. In the present study, we explore the adherence to such rules in the context of joining a small, freestanding conversational group of agents in VR. In particular, we focus on the behavior adopted by participants while walking towards the agents, and on whether ECAs were treated in the same way human agents normally are. 45 test subjects were invited by an ECA to walk towards the group by applying one of six possible politeness strategies; after freely joining the group, they were asked to rate the agent's politeness according to four distinct aspects (Clarity, Face loss, Positive face, and Negative face). Across all strategies, in 48% of the trials participants were successfully persuaded to join the group at an inconvenient location. Out of those trials, participants adhered to social conventions by not crossing the convex empty space between the group members (o-space) in 75% of them on average. Additionally, analysis of verbal and nonverbal behaviors in ECAs shows that direct request strategies are more effective than indirect ones, although in some cases they may be perceived as less polite.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **HCI theory, concepts and models**.

KEYWORDS

politeness theory, persuasiveness, virtual reality, small group behavior, embodied conversational agents

ACM Reference Format:
Alessandro Iop, Sahba Zojaji, and Christopher Peters. 2022. Don't Walk Between Us: Adherence to Social Conventions When Joining a Small Conversational Group of Agents. In *ACM International Conference on Intelligent Virtual Agents (IVA '22), September 6–9, 2022, Faro, Portugal.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3514197.3549676

IVA '22, September 6–9, 2022, Faro, Portugal

ACM ISBN 978-1-4503-9248-8/22/09.
https://doi.org/10.1145/3514197.3549676

1 INTRODUCTION

In the context of interaction between humans, a speaker's intention is translated through interaction to specific actions, utterances and expressions that convey meaning and affect both parties' subsequent approach in the ever-unfolding sequence of messages exchanged by them. Keeping one's behavior consistent with the intent and emotions behind it is defined as *face work* [8]. One way of persuading both oneself and others that the speaker is trustworthy, confident and, most importantly, non-threatening, is through gestures. Their shape, expressiveness and meaning are all contributing factors in the performed face work [10, 22]. Social behavioral norms, which are dictated by cultural and intrinsic principles, govern human conversations in all kinds of contexts; they are the protocol to which we adhere when performing face work as well as interpreting messages from a fellow *interactant*. We tend to adhere to these norms when communicating with *Embodied Conversational Agents* (ECAs) in the same way we do with other humans [2]. To this day, very little research has explored the implementation of such nonverbal behavior, alongside verbal utterances, in ECAs with the goal of investigating their *persuasiveness* when being interacted with by a real person. In particular, the effectiveness of specific behavioral strategies has not yet been thoroughly assessed when a persuasive intent is expressed by a virtual agent in a conversational context.

In order to investigate this issue, in the present study we consider two main questions:

(1) To what degree are participants persuaded by an ECA to join a small group at an *inconvenient location* according to six different *politeness strategies* (PSs), which are expressed by combinations of verbal and nonverbal behaviors? How does *perceived politeness* relate to the effectiveness of the PSs?
(2) For the cases in which participants are compliant with the agent's request to join the group at the inconvenient location, what are trajectories taken by participants to get there?

The study has been purposefully designed to present participants with a dilemma in which they must choose to either expend substantially more effort to join the group in a socially acceptable manner (i.e. by walking around the group), versus choosing a more economical route through the group members that violates social behavioral norms (i.e. by walking between the ECAs). This has been done in order to better quantify the degree to which participants will continue to engage in social "group-joining" behaviors with virtual agents even when they have not been asked to do so and are well aware that there is no real human presence behind them. Moreover, we are interested in determining whether test subjects' perception of the ECAs as proxies for the experimenter bears an impact on their walking behavior.

More specifically, in the present study participants were invited by a virtual ECA to join a small, free-standing conversational group by means of different combinations of verbal and nonverbal behaviors, which in turn implemented one of six possible PSs. For 18 total trials, each test subject joined the group by freely walking in the virtual room with their preferred trajectories. They were also aware that the ECAs were not controlled by the experimenter. *Virtual Reality* (VR) technologies were employed to provide participants with an immersive experience and a natural set of interaction and locomotion paradigms, as well as to assess whether results on perceived politeness hold irrespective of the *medium*, when compared to those from a more traditional desktop application.

We hypothesize that findings brought about in related work hold true irrespective of the medium employed, and our results will be in line with previous work in politeness theory and group theory. Conclusions presented here may impact future research endeavors by consolidating the foundations of existing work around the behavioral analysis of IVAs, especially in immersive environments. At the same time, it can provide designers with insights on how humans react to politeness behaviors shown by ECAs and how they perceive their persuasive intents, especially in the context of small conversational groups. A possible comparison between human-agent interactions in different types of media (desktop vs. VR), as well as between agents of different degrees of embodiment (human vs. robot vs. virtual) may then be partially informed by our findings.

2 RELATED WORK

Erving Goffman [8] defined *face* as "the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact", and consequently face work as "the actions taken by a person to make whatever he is doing consistent with face". In other words, the former can be viewed as the politeness of a person as perceived by themselves and others by means of that person's behavior, while the latter is the strategy that is employed with the aim of appearing polite. Face work is hence relying on the persuasiveness skills of an interactant, for which André et al. [2] argue nonverbal behaviors - such as gaze and gestures - are important when they convey importance, certainty, evaluation, benevolence, competence and emotion. Brown and Levinson [5] expanded on the concepts proposed by Goffman, by defining politeness as "an effort to mitigate or avoid doing action which damages an individual's public self-image", i.e. their face, which in turn can be perceived as positive or negative. They propose a categorization of possible strategies that can be employed to convey a message in a conversation, via *face-threatening acts*, according to their effect on the perception of politeness. These are: *not* performing a face-threatening act, *indirect* strategies (also referred to as off-record), *negative politeness* - based on the avoidance of imposition by the speaker, *positive politeness* - based on the avoidance of offensiveness, and *direct* strategies (also known as on-record).

The design of group formation and behaviors presented here was informed by several studies that delve into the topic of small group behavior in artificial systems. More specifically, a 2018 paper by Sai Krishna Pathi [21] explored the ideal position for agents to enable social interactions, while Althaus et al. [1] in 2004 focused on robot formations and movements in human-robot interaction, a more recent topic of discussion also faced by Jamy Li in 2015 [17] with regards to the feeling of presence. Two additional sources are worth mentioning in this context, namely Edward T. Hall's definition of four distinct group formation zones involving embodied social agents (*intimate space*, *personal space*, *social space* and *public space* [9]) and Adam Kendon's *F-formations* comprising the so called o-, p- and r- spaces [13]. While the former more broadly differentiates interactions according to the type of acquaintance between group members, the latter more specifically "describes the common space management in a group in which all members have equal, direct, and exclusive access to the group space". According to the *principle of least effort* [11], when presented with a choice between similarly rewarding options, people tend to prefer the option that requires less work or effort; on the other hand, people may be more inclined to prefer the option that requires more effort when it is considered to have a greater perceived value [12]. In the context of joining a conversational group, walking between the group members can therefore arguably be considered as requiring less work than walking around them, although previous research suggests that crossing the o-space cannot be considered as a conventional social behavior [7, 13, 14]. Social conformity and proxemics have been subjects of focus in recent studies around human-robot interaction, with results that at times support similarities to human-human interaction [18], and at times refute them [27, 28].

A study by Zojaji et al. [29] focused on the impact of different politeness behavior strategies on the trajectories taken by human participants when joining a virtual conversational group. In particular, the authors investigated in what measure five combinations of verbal and nonverbal behaviors (plus one control combination), modeled on Brown and Levinson's own strategies, influence the path taken by participants when invited by a virtual agent to join the conversational group via keyboard controls in the context of a simple application. Results from the study showed that "more direct and explicit politeness strategies have a higher level of success when requesting a participant to join a small group at an inconvenient location, but sometimes negatively impact their perception of the agent". Additionally, participants were more likely to be persuaded to join the group at inconvenient location and have a positive impression of the agent when a positive PS was applied. A further exploration of politeness strategies within the context of human-virtual agent interactions was subsequently presented by Terada et al. [26], who investigated the impact of different strategies in negotiations between participants and conversational virtual agents. Their study showed that adopting indirect and implicit expressions (namely, an *off-record* approach) enabled the extraction of greater concessions from the human participants in the negotiation, as opposed to a positive politeness approach. Other studies investigating human behavior when joining groups of conversational IVAs [6, 25] have not focused on quantifying the willingness of subjects to expend effort in order to do so "socially", and few have explored the free-standing small group interactions in immersive virtual environments (VEs) [20].

3 METHODS

3.1 Scenario

Two virtual conversational agents are placed at the center of a virtual room, facing each other and defining several locations of proximity, in relation to the participant's starting position. The agent facing the participant is defined as primary and is the one inviting the participant to join in, while the other is secondary. A virtual avatar, matching the declared gender of the participant, provides embodiment and aids in the orientation within the room. As depicted in Figure 1, participants could either choose to join in at a convenient position (in yellow) or at an inconvenient one (in red). In the former case, the least amount of objective effort (~7 steps) is required, whereas in the latter subjects could walk in a straight line across the o-space (~10 steps) or around the secondary ECA (~14 steps). These possible trajectories are depicted in yellow (the "convenient" route), red (the "unsocial" route) and green (the "inconvenient" route), respectively.

For each trial, test subjects start at a distance from the group to one side, and are asked to join the group by freely moving in the environment while keeping the following scenario in mind:

> *Three adults are in a language exchange meeting in an indoor environment. Two of them are intelligent virtual agents who are already in a group and a human (you) who is represented as an avatar. All of you have seen each other and already talked together for a while. You have left the group to read more about a topic. It took about 15 minutes to read about it. Now you are going to come back to the group and join them again to talk about your findings.*

To avoid biases, starting positions alternate between left and right (of the group) across trials. In addition to having complete freedom of movement, participants are reminded that the agent is completely virtual and non-existent in reality, so as to avoid interpreting the agent as a *proxy* of the experimenter. Six distinct PSs consisted of verbal and nonverbal behaviors: Baseline (BSL), Indirect (IND), Asking (ASK), Proposing (PRO), Commanding (CMD) and Pointing (PNT) (Table 1). These behaviors were implemented using Greta [19, 23], a virtual character engine that allows generating socio-emotional behaviors in order to build natural interactional scenarios with human users.

3.2 Structure of the study

The user study was structured as follows.

An online demographics survey collected information such as age, nationality, country of residency, proficiency in English, familiarity with Artificial Intelligence (AI) systems and VR, gender, handedness. Instructions were provided in written form on the structure and steps of the experiment, and clarifications were made by the experimenters upon request.

A "demo" scenario in VR, designed to make the participants accustomed to the VE, presented a group of ECAs with no behaviors. Subjects were able to explore the virtual room and observe the agents as long as they needed.

The main experiment in VR consisted of 18 trials (three for each of the PSs considered) performed consecutively and separated by short questionnaires. PSs were randomized for the first six trials, then repeated in the same order for the subsequent 12 trials. The between-trials questionnaires asked participants to what degree they agreed with four statements using a Likert scale, one for each of the politeness variables considered. Throughout the entire study, participants were asked to keep a specific scenario in mind, with the aim of evoking a realistic setting involving the group of conversational agents presented to them (see Figure 2).

An online debrief questionnaire[1], aimed at assessing participants' perception of the primary ECA as a *proxy* for the experimenter,

Table 1: Experiment PSs conditions and their related verbal and nonverbal behaviors. Note: in all conditions, the agent performed the same reaction (gaze at the participant and smile) prior to the politeness behaviors below. Nonverbal behavior labels are as following: open palm up (UP), open palm sideways (SIDE), open palm sideways and partly downwards (DOWN), pointing directly at a specific point with the index finger (POINT).

Condition	Verbal	Nonverbal
BSL	None	None
IND	"Welcome back!"	UP
ASK	"Would you like to come here?"	SIDE
PRO	"This place is waiting for you!"	DOWN
CMD	"Come here!"	POINT
PNT	None	POINT

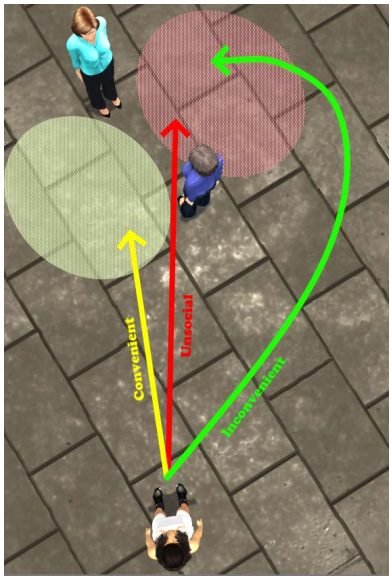

Figure 1: Top-down view of the positioning of the virtual agents (top) and the participant-controlled avatar (bottom), along with the three possible routes to join the group: "convenient" in yellow, "unsocial" in red, "inconvenient" in green.

[1] https://github.com/zojaji/zojaji.github.io/raw/master/assets/documents/IVA22-IopZojajiPeters-DebriefQuestionnaire.pdf

asked them whether in their minds they were supposed to follow the ECA's instructions or not, and what "following the instructions" meant to them. Before responding to the final section of the debrief questionnaire, asking participants to subjectively assess their effort in joining the nearest and farthest point of the virtual conversational group, they were asked to join the conversational group at both points while immersed in the "demo" scenario.

HTC VIVE Pro headsets and controllers were used in the present study. To enable participants to walk in the whole room, HTC VIVE wireless adapters were attached to the headset and charged by a wearable battery pack (see Figure 2). The controllers were used for interaction with User Interfaces (UIs) presenting, for instance, between-trials questionnaires; a virtual pointer was displayed on one of them, and the trigger button allowed selection/clicking of specific fields and buttons.

To allow subjects to comfortably walk around the ECAs in VR and to avoid accidental collisions with the walls and furniture of the real room (10x12 m^2 in size), Redirected Walking [24] and the freeze-turn method [4] were applied to the experimental design. Participants were asked to start the experiment at a particular location and orientation in the real room, before wearing the VR headset. For each trial, after joining the group of agents the between-trial questionnaire was presented in the VE at the starting position: they needed to turn around, walk back to the initial position and face the initial orientation in order to be able to interact with the questionnaire UI. At the beginning of the subsequent trial, they were therefore already at the correct position and location to begin anew. By slightly modifying the position of the virtual conversational group (to the left and to the right, alternately) across trials, participants kept the freedom of movement while being blind to the changes applied to the environment.

Figure 2: A participant immersed in the VE during the user study, standing at the prescribed starting position and orientation to begin each trial. Free walking in the room is enabled by the wireless adapter mounted on top of the HTC VIVE Pro headset, and connected to a battery pack.

3.3 Participants

The user study involved 45 participants with a good proficiency in English, recruited online and compensated with either a voucher or a cinema ticket. Of these, 9 were discarded due to external circumstances (technical problems, interruptions) or unnatural behavior - e.g. completing each trial too quickly, glossing over the between-trial questionnaires with automatism. Of the 36 participants considered in the analysis of results, aged between 18 and 39 (26±4.5), 36% (n = 13) were female and 64% (n = 23) were male. Additionally, 58% had basic (n = 14) or no (n = 7) previous experience with AI systems, the rest having intermediate (n = 11) or advanced (n = 4) experience in the field. Finally, concerning VR more participants had basic (n = 21) or no (n = 7) experience than those who had intermediate (n = 3) or advanced (n = 3) experience. Data collected from each participant was anonymized, and written consent was signed by them before the start of the experiment. Additionally, they were given the option to agree on having photos and videos taken during the experiment.

3.4 Metrics

The questions presented in the between-trials questionnaire during the user study were designed specifically to address four principal components of perceived politeness: *Clarity*, *Face loss* (i.e., offensiveness), *Positive face* (i.e., friendliness), *Negative face* (i.e., freedom of action). Questions were presented as brief sentences, and subjects were asked if they agreed through 7-point Likert scales. In addition, persuasiveness of the agent's request was assessed by whether each subject, during each trial, joined the conversational group at the inconvenient position. The frequency of o-space crossings - i.e. walking between the primary and secondary agents - was also recorded. From these two measures, rates for success in the ECA's request and for the adherence to social behavioral norms were computed. Finally, another measurement taken during the present study, but not considered during the analysis of results, was the movement paths taken by participants in each trial. The debrief questionnaire presented questions directly aimed at assessing: 1) whether participants viewed the primary ECA as a *proxy* for the experimenter; 2) the perceived subjective efforts in reaching a convenient and an inconvenient location in the conversational group while immersed in VR.

4 RESULTS

A between-subjects ANOVA test was performed to detect whether there were significant differences between participants who viewed the primary ECA as a *proxy* for the experimenter and those who did not, with persuasiveness/politeness variables as dependent and their response to the relevant debrief-questionnaire item as independent. No significant differences were found between these results and those obtained over the entire population.

4.1 Persuasiveness

A within-subjects, repeated measures ANOVA test with *Condition*[2] (with six levels) as within factor and persuasiveness as dependent

[2]For the purpose of the present study, the terms "strategy" and "condition" are to be considered equivalent.

Table 2: Breakdown of *success rate* and of *o-space crossings rate*. Across all conditions, in 48% of the trials participants were successfully persuaded to join the group to the farthest position. Of those cases, participants adhered to social conventions by not crossing the o-space in 75% of the trials on average.

Condition	BSL	IND	ASK	PRO	CMD	PNT
Requested	106	101	103	104	99	102
Successful	21	19	55	66	71	63
Success rate	20%	19%	53%	63%	72%	62%
O-space crossings	5	8	15	15	17	15
Adherence to social conventions	76%	58%	73%	77%	76%	76%

variable was conducted on the data collected from the study. Overall, there was a statistically significant effect of *Condition* on joining the group at the inconvenient location, $F(3.730, 130.553) = 26.480$, $p < .001$. As shown in Figure 3, a Bonferroni post-hoc comparison suggests that direct conditions are more persuasive, with a significant difference between the conditions BSL, IND and the remainder of the conditions (ASK, PRO, CMD, PNT). Table 2 shows the success rate of the primary ECA for each PS, i.e. the overall ratio of trials in which participants joined the group at the inconvenient location. A clear distinction can be made between indirect and direct strategies, with the latter being more effective in their request than the former. Across all conditions, 48% of participants joined the group to the inconvenient position.

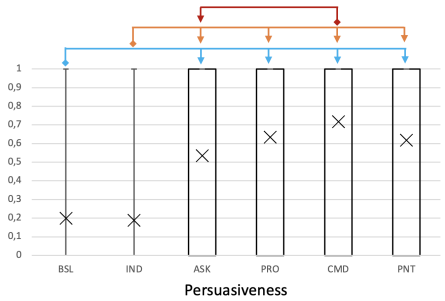

Figure 3: *Persuasiveness*. Effect of politeness strategy on the participant joining at the farthest location. The Y-axis plots the distribution over all the participants and trials of the persuasiveness boolean value. Higher values indicate a stronger persuasiveness in the strategy. In all boxplots, significance at .05 is indicated by arrows connecting different PSs.

To support these findings that group indirect and direct conditions separately, the number of o-space crossings for each PS was averaged across conditions belonging to each group and analyzed in the same way; a significant difference was found between the number of times participants walked across the o-space for indirect and direct conditions, $F(1,35) = 15.326$, $p < .001$. Similarly, effort ratings in joining the group to the closest and farthest positions reported by participants in the debrief questionnaire were analyzed. A within-subjects, repeated measures ANOVA with *Condition* group (direct and indirect) as within factor and *Effort* (near and far) as dependent variables was performed; a significant difference in effort was found between joining the closest and farthest position, $F(1, 35) = 41.842$, $p < .001$.

Table 3: Breakdown of the number of trials per block in which the agent requested and successfully persuaded participants to join the group at the farthest location, and of the related number of trials in which participants walked between agents.

Block (trials)	I (1-6)	II (7-12)	III (13-18)	Total
Requested	206	205	204	615
Successful	105	96	94	295
Success rate	51%	47%	46%	48%
O-space crossings	26	30	19	75
Adherence to social conventions	75%	69%	80%	75%

4.2 Adherence to social behavioral norms

Table 2 also shows to what degree participants who were successfully persuaded to join the group at the inconvenient location adhered to social behavioral norms. In other words, the percentage of those test subjects who behaved as they would normally do with other humans by deliberately choosing *not* to cross the o-space. Across all conditions, 75% of successful requests resulted in participants expending effort to travel the most inconvenient route (around the group). Additionally, with the exception of the IND strategy, this indicator is always above 70%.

To test possible effects of fatigue, expectations and priming on test subjects throughout the experiment, the adherence to social behavioral norms was compared across three trial blocks, which divide all 18 trials into groups of six. As shown in Table 3, there is no significant difference in the way participants joined the group at the farthest location from the first trials to the last ones. In other words, they consistently chose to obey to the conventional human interaction principles, even after the different PSs were presented to them multiple times.

4.3 Politeness

A within-subjects, repeated measures ANOVA test with *Condition* (with six levels) as within factor and *Clarity*, *Face loss*, *Positive face* and *Negative face* as dependent variables was conducted on the data collected from the study. Overall, there were statistically significant effect of all these factors on the perception of politeness in the agent's request. Bonferroni post-hoc comparisons show that:

- BSL was perceived as least clear among all strategies and PNT was less clear than all other direct conditions (see Figure 4a);

- CMD and PNT were perceived as most offensive, while IND and ASK as least offensive (see Figure 4b);
- IND, ASK and PRO were perceived as most friendly, while BSL as least friendly (see Figure 4c);
- CMD and PNT made participants feel the most constrained in terms of their freedom of action (see Figure 4d);

5 DISCUSSION

Our primary hypotheses for the present study were that conclusions from related work would be confirmed irrespective of the medium employed, and that our results would be in line with the group and politeness theories that were formally introduced in past literature.

Results from the present study show that the verbal and nonverbal behaviors adopted by the ECAs to implement the PSs appeared to be understood by test subjects, and some of these were effective in persuading them to join the conversational group at an inconvenient location. Analysis of the primary agent's politeness, as perceived by participants, show that conclusions proposed by Zojaji et al. [29] hold true despite the fact that VR provides a more natural interaction and locomotion, as well as a more immersive experience, when compared to a more traditional Desktop application. This is suggested in first place by success rates measured for each PS (see Table 1), which are similar to those measured for a Desktop application. Secondly, statistical analysis performed on the four indicators of perceived politeness yield results comparable to the corresponding ones presented in the previous research.

In particular, the absence of any verbal and nonverbal behavior is associated with low clarity, an observation that is in line with Brown and Levinson's categorization of strategies used to avoid face-threatening situations [5], as well as with the theory proposed by André et al. [2] around the importance of nonverbal behavior in conveying persuasiveness. More imposing and direct strategies, despite being somewhat clear in their intention, are perceived as more offensive, less friendly and inhibiting the interactant's freedom of action, ultimately leading to the agent's Face loss; definitions made by André et al., who propose that "one who wishes to persuade must appear credible, trustworthy and confident", are confirmed.

We can suggest a ranking of politeness strategies based on their effectiveness in the context of joining a small conversational group of ECAs, categorizing them from best to mid to worst and assigning them a score of three, two and one, respectively. Their scores, for each variable under consideration, are therefore assigned according to their associated results and relationships of statistical significance to one another. By summing up all scores assigned to each PS, we can draw a conclusion on which one has the best success in the specific scenario we presented. As shown in Table 4, in both the VR and Desktop applications, Asking (ASK) is the overall best strategy, while Baseline (BSL) is the worst. In other words, a direct but not imposing strategy achieves the highest success in persuading a human interactant and the ECA is perceived to be most polite while performing it. On the other hand, a strategy that does not employ any form of interaction evokes impoliteness and does not convince the human interactant to join the group. This observation, from the point of view presented by Brown and Levinson [5], indicates that negative politeness strategy is the best way in this context that the agent can avoid loss of face.

Table 4: Ranking of PSs from best (top) to worst (bottom), with a comparison between the previous and current studies. Total cumulative scores are listed for each PS. Baseline (BSL) is overall the worst while Asking (ASK) is the best.

Rank	Strategy	Score VR	Score Desktop
Best	ASK	14	13
	PRO	12	12
	IND	12	11
	CMD	8	10
	PNT	7	8
Worst	BSL	7	7

5.1 Persuasiveness

In assessing the degree to which participants are persuaded to join the conversational group at an inconvenient location, it appears that indirect strategies are less effective than direct ones: there is a statistically significant difference between PSs that use indirect language and gestures or none at all to express an interactant's need and those who do not. In the case of BSL in particular, a non-zero success rate may be due to its low perceived clarity as well as it being possibly preceded by other PSs in the experiment, which in turn would create expectations (*priming* effect [3]) in participants' minds; an issue that could be further investigated in future studies. Utterances and nonverbal behaviors that convey more certainty, importance and emotion seem therefore to have a bigger impact on persuasion; such observation is in line with the theory proposed by André et al. [2], as well as with conclusions drawn by Zojaji et al. [29]. This comparison appears however to refute the conclusions drawn by Terada et al. [26], who claim that off-record strategies are more effective than Positive face strategies in getting concessions from the human interactant.

The absence of any significant difference between participants who viewed the agent as a proxy for the experimenter and those who did not rules out the possibility that in joining the group of ECAs they behaved in the way they thought they were supposed to. In other words, test subjects *deliberately* chose to expend more effort by walking to an inconvenient location despite their choice was not sanctioned and no feedback was given them by the experimenter during the experiment. Although 23 out of 36 test subjects indicated that in their minds they were supposed to follow the primary ECA's instructions, further elaboration on what this meant to them was inconsistent and at times incorrect. This was the case when participants identified the agent's suggestions with the task provided by the instructions (i.e. to "join the group"), or when they contradicted themselves in their answers. The absence of a proxy effect on test subjects is also supported by the lack of any feedback from the ECA on the trajectory walked by participants when joining the group.

5.2 Adherence to social behavioral norms

Of all successful trials for the six PSs considered, the vast majority (75% in total) show adherence by participants to social behavioral norms. When joining the group to an inconvenient locations, participants mostly behaved in a social manner towards the ECAs

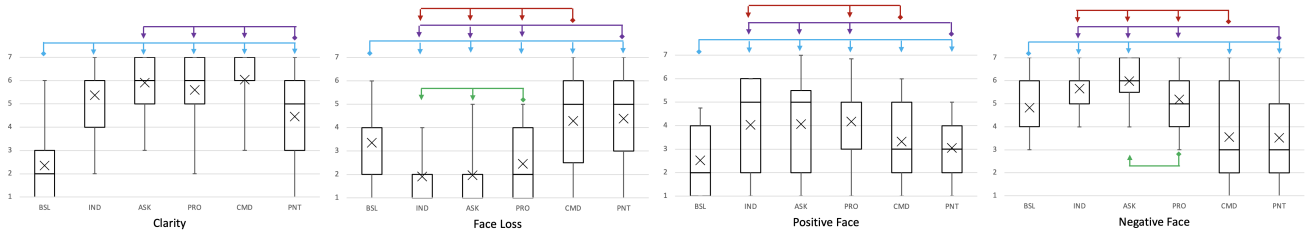

Figure 4: *Perception of politeness.* Effect of politeness strategy in terms of: a) Clarity; b) Face loss; c) Positive face; d) Negative face. Significance at 0.05 is indicated by the arrows on top connecting different PSs.

and decided to take the longer route around the group to join them, therefore not crossing the o-space. Additionally, despite being aware that the agents were not human like them, they never walked through or collided with the 3D models of the ECAs. This indicates that, in this context, people obey to conventional interaction norms with ECAs in VR, even when substantially more effort is required to do so. Such observation is corroborated by the comparison between trial blocks presented in Table 3: the adherence to social behavioral norms *increased* with the progress of the experiment, regardless of any possible effects of fatigue, priming, boredom and "proxying".

As previously mentioned, humans usually tend to minimize the expended effort in social interactions unless doing so is not considered to be rewarding. It appears that, in the present study, participants valued respecting the space between ECAs and accepting their request - thus behaving politely - more than saving effort by walking to a convenient place in the virtual room. Surprisingly, they kept doing so throughout the experiment, even after all PSs were repeatedly presented to them and even knowing that no reward or punishment was given for their choice of trajectory. The interaction with the agents ended in each trial with reaching the final joining position and was not protracted further; nevertheless, test subjects consistently decided to walk the longer route in the vast majority of successful trials.

5.3 Limitations and future work

In the present work, none of the data on the specific trajectory taken by participants in space, collected during the user study, has been analyzed. Future work can delve deeper into this aspect by considering the overall distance covered by test subjects when walking in the room, as well as the time they took to join the conversational groups and any particular deviation from a more convenient approach to the target locations. This can in turn solidify results presented here and at the same time yield an insightful analysis on participants' *social behaviors*. Additionally, a comparison with previous work by Zojaji et al. [29] can be made by analyzing the same data collected in their own experiment. Moreover, a user study delving deeper into the perception of ECAs as proxies for the experimenter can be designed, with the goal of assessing whether subjects carry expectations and biases related to the instructions given to them before the experiment.

Through the four politeness variables considered in our study, we have discussed how PSs, i.e. the agent's face work, influence its perceived politeness and consequently the routes (including o-space crossings) taken by participants. An obvious continuation along this line would explore interaction with robots and/or other humans. Research conducted by Vollmer et al. [28] and Neggers et al. [18] in human-robot interaction, could, together with the present study and more general work on social proxemics in IVAs [6, 15, 16, 20], present a starting point for an insightful comparison between multiple kinds of agent in the same context.

6 CONCLUSIONS

This study addresses the persuasion of human interactants in joining a small, free-standing group of ECAs and their adherence to conventional interaction norms that are traditionally applied in human-human conversations. Results show that a clear distinction can be made between indirect and direct politeness strategies when it comes to the walking behavior shown by test subjects, along with their perception of the inviting agent's politeness. The absence of verbal and nonverbal behaviors results in low perceived clarity, low friendliness, high offensiveness and participants walking a shorter distance in space. Meanwhile, more imposing and direct strategies are more persuasive despite being perceived as more offensive and less friendly, and result in a lower freedom of action. We conclude that a direct, but not imposing, strategy is the best approach in this context.

Additionally, results show that people tend to adhere to social behavioral norms when joining the conversational group, despite being aware that no sanction is given for their choice of trajectory and that agents are not controlled by other humans. The vast majority of successful requests, in fact, resulted in participants not crossing the o-space or the 3D models of any ECA while walking in the virtual environment. We therefore suggest that people expend effort and behave socially when interacting with small groups of conversational agents.

ACKNOWLEDGMENTS

We wish to thank Catherine Pelachaud for her help in the design of the original study and the reviewers for their constructive suggestions that helped to improve the paper. This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 765955.

REFERENCES

[1] Philipp Althaus, Hiroshi Ishiguro, Takayuki Kanda, Takahiro Miyashita, and Henrik Iskov Christensen. 2004. Navigation for human-robot interaction tasks. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004.* IEEE, 1894–1900 Vol.2. https://doi.org/10.1109/ROBOT.2004.1308100

[2] Elisabeth André, Elisabetta Bevacqua, Dirk Heylen, Radoslaw Niewiadomski, Catherine Pelachaud, Christopher Peters, Isabella Poggi, and Matthias Rehm. 2011. *Non-verbal Persuasion and Communication in an Affective Agent*. Springer Berlin Heidelberg, 585–608. https://doi.org/10.1007/978-3-642-15184-2_30

[3] John A Bargh and Tanya L Chartrand. 2014. The mind in the middle: A practical guide to priming and automaticity research. (2014).

[4] Evren Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Rajiv Dubey. 2019. Locomotion in virtual reality for room scale tracked areas. *International Journal of Human-Computer Studies* 122 (Feb 2019), 38–49. https://doi.org/10.1016/j.ijhcs.2018.08.002

[5] Penelope Brown and Stephen C. Levinson. 1978. *Universals in language usage: Politeness phenomena* (3 ed.). Cambridge University Press, 56–311.

[6] 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 Transactions on Interactive Intelligent Systems* 6, 2 (Aug 2016), 1–33. https://doi.org/10.1145/2914796

[7] T Matthew Ciolek and Adam Kendon. 1980. Environment and the spatial arrangement of conversational encounters. *Sociological Inquiry* 50, 3-4 (1980), 237–271.

[8] Erving Goffman. 1967. *On Face Work*. 5–45.

[9] Edward T. Hall. 1990. *The hidden dimension*. Anchor Books.

[10] Björn Hartmann, Maurizio Mancini, and Catherine Pelachaud. 2002. Formational parameters and adaptive prototype instantiation for MPEG-4 compliant gesture synthesis. In *Proceedings of Computer Animation 2002 (CA 2002)*. IEEE, 111–119.

[11] CA Hull. 1943. Behaviour System and Principles of Behaviour.

[12] Michael Inzlicht, Amitai Shenhav, and Christopher Y Olivola. 2018. The effort paradox: Effort is both costly and valued. *Trends in cognitive sciences* 22, 4 (2018), 337–349.

[13] Adam Kendon. 1990. *Conducting interaction: patterns of behavior in focused encounters*. Cambridge University Press.

[14] Adam Kendon. 1992. The negotiation of context in face-to-face interaction. Goodwin C. & Duranti, A.(eds) Rethinking context: Language as an interactive phenomenon.

[15] Peter Khooshabeh, Sudeep Gandhe, Cade McCall, Jonathan Gratch, Jim Blascovich, and David Traum. 2011. *The Effects of Virtual Agent Humor and Gaze Behavior on Human-Virtual Agent Proxemics*. Lecture Notes in Computer Science, Vol. 6895. Springer Berlin Heidelberg, Berlin, Heidelberg, 458–459. https://doi.org/10.1007/978-3-642-23974-8_61

[16] Jan Kolkmeier, Jered Vroon, and Dirk Heylen. 2016. *Interacting with Virtual Agents in Shared Space: Single and Joint Effects of Gaze and Proxemics*. Lecture Notes in Computer Science, Vol. 10011. Springer International Publishing, Cham, 1–14. https://doi.org/10.1007/978-3-319-47665-0_1

[17] Jamy Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies* 77 (May 2015), 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

[18] Margot M. E. Neggers, Raymond H. Cuijpers, Peter A. M. Ruijten, and Wijnand A. IJsselsteijn. 2022. Determining Shape and Size of Personal Space of a Human when Passed by a Robot. *International Journal of Social Robotics* 14, 2 (Mar 2022), 561–572. https://doi.org/10.1007/s12369-021-00805-6

[19] Radoslaw Niewiadomski, Elisabetta Bevacqua, Maurizio Mancini, and Catherine Pelachaud. 2009. Greta: an interactive expressive ECA system. In *AAMAS '09: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*, Vol. 2. International Foundation for Autonomous Agents and Multiagent Systems, 1399–1400. https://dl.acm.org/doi/abs/10.5555/1558109.1558314

[20] David Novick and Aaron E. Rodriguez. 2021. *A Comparative Study of Conversational Proxemics for Virtual Agents*. Lecture Notes in Computer Science, Vol. 12770. Springer International Publishing, Cham, 96–105. https://doi.org/10.1007/978-3-030-77599-5_8

[21] Sai Krishna Pathi. 2018. Join the Group Formations using Social Cues in Social Robots. Association for Computing Machinery (ACM), 1766–1767. https://www.diva-portal.org/smash/record.jsf?dswid=-7816&pid=diva2%3A1283242

[22] Isabella Poggi and Catherine Pelachaud. 2008. Persuasion and the expressivity of gestures in humans and machines. *Embodied communication in humans and machines* (2008), 391–424.

[23] Isabella Poggi, Catherine Pelachaud, F De Rosis, V Carofiglio, and B De Carolis. 2006. *Greta. A believable embodied conversational agent*. Springer Science & Business Media, 3–25.

[24] Sharif Razzaque, Zachariah Kohn, and Mary C. Whitton. 2001. Redirected Walking. *Eurographics 2001 - Short Presentations* (2001). https://doi.org/10.2312/EGS.20011036

[25] Matthias Rehm, Elisabeth André, and Michael Nischt. 2005. *Let's Come Together — Social Navigation Behaviors of Virtual and Real Humans*. Lecture Notes in Computer Science, Vol. 3814. Springer Berlin Heidelberg, Berlin, Heidelberg, 124–133. https://doi.org/10.1007/11590323_13

[26] Kazunori Terada, Mitsuki Okazoe, and Jonathan Gratch. 2021. Effect of politeness strategies in dialogue on negotiation outcomes. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 195–202.

[27] Anna-Lisa Vollmer, Katrin Solveig Lohan, Kerstin Fischer, Yukie Nagai, Karola Pitsch, Jannik Fritsch, Katharina J. Rohlfing, and Britta Wredek. 2009. People modify their tutoring behavior in robot-directed interaction for action learning. In *2009 IEEE 8th International Conference on Development and Learning*. IEEE, Shanghai, China, 1–6. https://doi.org/10.1109/DEVLRN.2009.5175516

[28] Anna-Lisa Vollmer, Robin Read, Dries Trippas, and Tony Belpaeme. 2018. Children conform, adults resist: A robot group induced peer pressure on normative social conformity. *Science Robotics* 3, 21 (Aug 2018), eaat7111. https://doi.org/10.1126/scirobotics.aat7111

[29] Sahba Zojaji, Christopher Peters, and Catherine Pelachaud. 2020. Influence of virtual agent politeness behaviors on how users join small conversational groups. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. ACM, 1–8. https://doi.org/10.1145/3383652.3423917