



A Comparative Study of Conversational Proxemics for Virtual Agents

David Novick^(✉) and Aaron E. Rodriguez

The University of Texas at El Paso, El Paso, TX 79968, USA
novick@utep.edu, aerodriguez14@miners.utep.edu

Abstract. This paper explores proxemics—interpersonal distances—in conversations with virtual agents in virtual reality. While the real-world proxemics of human-human interaction have been well studied, the virtual-world proxemics of human-agent interaction are less well understood. We review research related to proxemics in virtual reality, noting that the previous research has not addressed proxemics with actual conversation, describe an empirical methodology for addressing our research questions, and present our results. The study used a repeated-measures, within-subjects design and had 23 participants. In the study, participants approached and conversed with a virtual agent in three conditions: no crowd, small crowd, and large crowd. The participant’s distance from the agent with whom they conversed was recorded at 60 frames/second by VAIF’s proxemics tool. Our results suggest that humans in a virtual world tend to position themselves closer to virtual agents than they would relative to humans in the physical world. However, the presence of other virtual agents did not appear to cause participants to change their proxemics.

Keywords: Embodied conversational agents · Initiative management · Multiparty interaction

1 Introduction

This paper explores proxemics—interpersonal distances—in conversations with virtual agents in virtual reality. While the real-world proxemics of human-human interaction have been well studied, the virtual-world proxemics of human-agent interaction are less well understood. Do dyadic conversational proxemics in virtual reality resemble dyadic conversational proxemics in the physical world? Does the presence of other virtual agents in the environment affect people’s proxemic behaviors? The answers to these questions can inform the development of virtual-reality applications with human-agent interactions that feel more realistic and more fluid. For example, a model of how an agent should react when a human enters the room [1] could be augmented with better timing. And any virtual-reality application where an agent initiates a conversation when a human approaches (e.g., [2, 3]) could be improved through timing of the initiation that provides human-agent proxemics at a socially appropriate distance. Initiating the conversation too early would lead to a greater-than-appropriate distance, likely leading

the human conversant to think the agent is, literally, standoffish. And initiating the conversation too late would lead to a less-than-appropriate distance, likely leading the human conversant to feel uncomfortable about the closeness. Developers of human-agent interaction applications in virtual-reality application face these issues on a practical basis. For example, the virtual-reality application shown in Fig. 1 involves multiple conversations with different agents, plus dozens of non-speaking agents in the scene.



Fig. 1. User’s view of a scene in [1], in which the user is conversing with the foreground character. In this application, the user converses with seven different agents, in both indoor and outdoor settings. The presence of other agents in the scene may provide the user with guidance on appropriate proxemics. Reproduced with permission.

To address these issues, we review research related to proxemics in virtual reality, noting that the previous research has not addressed proxemics with actual conversation, describe an empirical methodology for addressing our research questions, and present our results. Our study suggests that humans in a virtual world tend to position themselves closer to virtual agents than they would relative to humans in the physical world. However, the presence of other virtual agents did not appear to cause participants to change their proxemics.

2 Review of Related Research

The research literature for proxemics among humans is extensive (e.g., [4–6]); see [7] for a review of the foundational work. In general, observational research has indicated that people have preferred interpersonal distances for conversation and that these distances vary as function of culture and of the relationship between the conversants. According to Hall et al., [4], typical distances for conversations (“personal space”) range from 0.46 to 1.2 m. This range is roughly consistent with the cultural variation observed by Herrera, Novick, Jan, and Traum [6], who observed mean interpersonal distances in dyadic conversation to vary across cultures from 1.09 to 1.66 m.

With respect to proxemics in virtual rather than real worlds, the literature is more limited. At least two studies have conducted observational experiments in which participants

judge the appropriateness of proxemics of groups of virtual agents. These studies sought to determine perceptually plausible proxemics in virtual worlds for cultural training [8] (see Fig. 2) and for video gaming [9].



Fig. 2. Example animation of North American proxemics model used as stimulus in [5]. Reproduced with permission.

But do humans carry over their proxemics behaviors and values from the real world to virtual worlds? Other studies have addressed proxemics more directly, with first-person interaction by human participants with virtual agents. These studies were conducted for various reasons (e.g., male vs. female, screen vs. immersive) and reported varying results. Where the human and the agent had mutual gaze (or similar), results ranged from 0.4 m [10] to 1.76 m [11]. Some studies (e.g., [12]) tended to confirm that virtual proxemics were similar to real-world proxemics, while others did not. Table 1 summarizes these findings. We note that there are inconsistencies across the reported results, with some studies reporting mean minimum distance and others mean distance.

Table 1. Human-agent proxemics results.

Study	Result (meters)	Notes
Bailenson et al., 2001	0.40	Mean minimum distance
Bailenson et al., 2003	0.54	Male mean minimum distance
Iachini et al., 2014	0.61	Mean stopping distance
Zibrek et al., 2017	1.76	Mean distance
Peters et al., 2018	1.23	Mean stopping distance

In some studies, the human walked up to (and sometimes around) the agent. In others, the agent walked toward the human. But in all of these studies, though, the human and the agent did not interact conversationally. Thus, here we address the open question of whether virtual-world proxemics, in the context of actual conversation, resemble real-world proxemics. Additionally, at least one of the observational proxemics studies [8] reflects an assumption that people will adapt their proxemics behaviors to the proxemics context. That is, people will stand more closely to their conversational partner if there are more people around and these people are closer together. Is this true? Thus, this paper addresses two related issues: The correspondence of virtual to real-world proxemics for conversational interaction and the effects of the whether other agents are nearby.

3 Methodology

To address these issues, we formed three hypotheses:

1. The mean distance between the participant and the ECA with which they are conversing, when only the ECA is present, will be the same as the mean distance between human conversants in a social context.
2. The standard error of the distance between the participant and the ECA with which they are conversing, when only the ECA is present, will be the same as the standard error of the distance between human conversants in a social context.
3. Participants will stand closer to the ECA with which they are conversing when the number of nearby ECAs increases.

To test the hypotheses, we conducted a within-subjects experiment with three conditions:

1. The participant and the virtual agent are alone.
2. The participant and the virtual agent have a few agents, with relatively distant proxemics, nearby.
3. The participant and the virtual agent have more agents, with relatively closer proxemics, nearby.

Twenty-three participants participated in the study, all of whom were undergraduate students at a public R1 university. Seventeen of the participants were male and five were female. Two of the participants had previously interacted with an embodied conversational agent. In the study, participants interacted with a female agent, and any surrounding agents, in a virtual world created with VAIF [13] and Unity, viewed with an HTC Vive headset. VAIF is both an authoring system for human-ECA interaction and a run-time system for executing these interactions, all built within Unity. The human-agent interactions in this study were entirely automated.

An earlier proxemics study with avatars [14] used a virtual world in which there were both boundaries and room to move (see Fig. 3), and the virtual world we used for this study used a similar layout (see Fig. 4).

Following our research protocol, we briefed the participants on the consent form and asked to read and sign the form, the participants completed a pre-interaction questionnaire on demographic information and then used the HTC Vive headset with a



Fig. 3. Virtual world used by Cafaro et al. [14] in an avatar-based study of proxemics. Reproduced with permission.

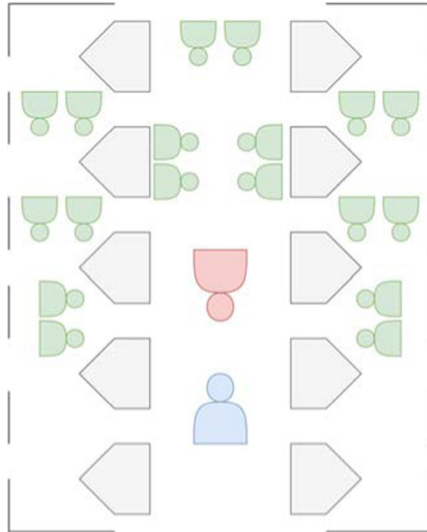


Fig. 4. Layout of virtual world for conversational proxemics study. The agent, shown in red, remains stationary. The human, shown in blue, enters from the bottom of the layout and walks toward the agent. Other agents, shown in green and non-speaking, represent the “large crowd” condition.

microphone headset to participate in three trials, balanced for order across subjects, corresponding to the experimental conditions listed above. Each trial had a different conversation, also balanced for order across subjects and experimental condition. The conversations, which were fully automated using VAIF, dealt with food, movies, and vacations. For example, Fig. 5 presents the beginning of the conversation about movies.

Agent: I love watching movies in my free time. Have you seen any movies lately?
 User: Yes, I think so, of course?
 Agent: That's awesome.
 (OR) No, I don't think so, no I don't, I doubt it, ...
 Agent: It's okay.
 Agent: I was planning on having a movie night later today. But I do not know which movie to watch. Are there any movies you would recommend?
 User: [wildcard]
 Agent: I have heard that the new Spiderman movie is really good! Have you heard about it before?
 User: Yes, I have, ...
 Agent: Oh cool!
 (OR) No, I have not, not really
 Agent: Oh ok.
 Agent: I saw the first Spiderman movie in the Galaxy theatre. I love watching movies at the theatre because I get to eat popcorn and drink icees. Where do you like watching movies?

Fig. 5. Beginning of conversation about movies. The *wildcard* notation indicates that any utterance from the user is accepted.

Each participant's distance from the agent with whom they were conversing was continuously recorded at 60 frames/second by VAIF's proxemics tool [15], measured between the center of the agent's head and the center of the participant's head. We confirmed the accuracy of the proxemic tool's measurements by physically verifying that the reported distance matched the actual distance; the largest disparity was 2 cm in a one-meter distance. The agent was triggered to begin speaking one second after the participant came near the agent. The trigger-distance varied as a function of the condition: 2 m in the no-crowd condition, 1.5 m in the small-crowd condition, and 1 m in the large-crowd condition. However, in all cases, the participants stopped moving before the agent began to speak, which means that the trigger distances did not affect the study's results.

For analysis purposes, we measured both minimum distance, to enable comparison with the results in [10] and [16], and mean distance. In the case of mean distance, we calculated this for the each third of each participant's interactions, and for our results we look at the middle third. Figures 6, 7, and 8 show representative images of the participant's view in the no-crowd, small-crowd, and large-crowd conditions, respectively.



Fig. 6. Participant's view in the no-crowd condition



Fig. 7. Participant's view in the small-crowd condition



Fig. 8. Participant's view in the large-crowd condition

4 Results

This study examined two main issues: the correspondence of virtual to real-world proxemics for conversational interaction and the effects of having other agents nearby. Our results suggest that people interacting conversationally with agents in virtual worlds exhibit proxemics that are much closer than those for real-world interaction but that the presence of other agents nearby has little or no effect on people’s conversational proxemics.

Our first hypothesis was that the mean distance between the participant and the ECA with which they are conversing, when only the ECA is present, would be the same as the mean distance between human conversants in a social context. The proxemics for dyadic conversations of speakers of American English, reported in [6] had a mean of 1.66 m and a standard error of 0.05. In contrast, the participants in our study produced proxemics values that were much closer, as indicated in Table 2. These values are roughly in line with those in the study report in [17], which involved non-conversational interaction. The values appear larger than those reported in [10], which we expect is due to two factors. First, as with the other previous studies, the interaction in [10] was non-conversational. Second, and probably more significant, the participants’ tasks in [10] involved reading text—the agent’s name and a number—from patches on the front and back, respectively of the agent. This meant that the participant, as a result of the task, would necessarily have to approach closely to the agent. This effect was likely enhanced by the agent’s standing still, rather than interacting, even physically, with the participant. In short, conversational proxemics in virtual worlds appear to be closer than in the real world, but not as close as non-conversational proxemics in the virtual world.

Table 2. Mean average proxemics in meters, by experimental condition and time divisions.

Condition	No crowd	Small crowd	Large crowd
Mean (1st third)	0.810	0.817	0.827
Mean (2d third)	0.835	0.821	0.854
Mean (3d third)	0.814	0.847	0.847

Our second hypothesis was that the standard error of the distance between the participant and the ECA with which they are conversing, when only the ECA is present, would be the same as the standard error of the distance between human conversants in a social context. For the mean dyadic conversational proxemics reported in [3] of 1.66 m, the standard error was 0.05. In our study, for the minimum proxemics (following Bailenson et al., 2001), the mean was 0.713 m, and the standard error was 0.042. This suggests that the range of variation in proxemics among the participants in the virtual conversations in our study is roughly the same as the range of variation in proxemics in the physical face-to-face conversation in [6].

Our third hypothesis was that participants will stand closer to the ECA with which they are conversing when the number of nearby ECAs increases. We expected that

both the presence of additional agents and these agents' modeling of proxemics would affect the participant's behavior. However, our results suggest that this hypothesis is not true, at least in the experimental conditions of this study. Looking at the case where the difference in proxemics should be most pronounced, no crowd vs. large crowd, the mean minimum proxemics are 0.713 (stdev 0.20) and 0.695 (stdev 0.23) m, respectively. This difference is not significant (two-tailed paired t-test: $p > 0.45$). Cohen's d was 0.08435, indicating that the effect size is between small and very small. It is unlikely that the lack of statistical significance was due to a small sample size; given these results, even for $N = 400$ the statistical power would be only 20%.

5 Conclusion

In this paper, we reviewed research related to proxemics in virtual reality. We observed that prior studies of human-agent proxemics did not involve conversational interaction. Accordingly, we designed and conducted a study to examine correspondence of virtual to real-world proxemics for conversational interaction and the effects of whether other agents are nearby. In the study, participants approached and conversed with a virtual agent in three conditions: no crowd, small crowd, and large crowd. Our results suggest that humans in a virtual world tend to position themselves closer to virtual agents than they would relative to humans in the physical world. However, the presence of other virtual agents did not appear to cause participants to change their proxemics.

This study is subject to a number of possible limitations, although it appears that a low N is not one of these, as discussed above. Rather, the first limitation is that the "crowd" agents may not have been sufficiently salient due to their distance from the conversing agent and, possibly, due to the structures in the environment. Second, there may be familiarization effects across the three trials that we did not detect. Consequently, we plan future work with (1) a set-up more faithfully resembling that in [14] so that the area has only external walls rather than interior structures and (2) "crowd" agents that are more numerous and closer to the conversing agent. Indeed, this point could inspire future work determining the circumstances (number of agents, proximity to the participant) in which crowd size and proxemics begin to affect the participant's own proxemics.

References

1. Novick, D., et al.: The Boston Massacre history experience. In: Proceedings of 19th ACM International Conference on Multimodal Interaction, 13–17 November 2017, Glasgow, Scotland, pp. 499–500 (2017)
2. Novick, D., Hinojos, L.J., Rodriguez, A.E., Camacho, A., Afravi, M.: Conversational interaction with multiple agents initiated via proxemics and gaze. In: 6th Annual International Conference on Human-Agent Interaction (HAI 2018), 15–18 December 2018, Southampton, UK (2018)
3. Novick, D., Adoneth, G., Manuel, D., Grís, I.: When the conversation starts: an empirical analysis. In: Joint Proceedings of the IVA 2012 Workshops. Otto von Guericke University Magdeburg, Santa Cruz, USA, pp. 67–74 (2012)
4. Hall, E.T., Birdwhistell, R.L., Bock, B., Bohannon, P., Diebold, A.R., Jr., et al.: Proxemics [and comments and replies]. *Curr. Anthropol.* **9**(2/3), 83–108 (1968)

5. Kendon, A.: Spacing and orientation in co-present interaction. In: Esposito, A., Campbell, N., Vogel, C., Hussain, A., Nijholt, A. (eds.) *Development of Multimodal Interfaces: Active Listening and Synchrony*. LNCS, vol. 5967, pp. 1–15. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12397-9_1
6. Herrera, D., Novick, D., Jan, D., Traum, D.: Dialog behaviors across culture and group size. In: Stephanidis, C. (ed.) *UAHCI 2011*. LNCS, vol. 6766, pp. 450–459. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21663-3_48
7. Altmann, I., Vinsel, A.M., Altmann, J., Wohlwill, J.F.: Personal space, an analysis of ET Hall's proxemic framework. *Hum. Behav. Environ.*, 181–259 (1977)
8. Jan, D., Herrera, D., Martinovski, B., Novick, D., Traum, D.: A computational model of culture-specific conversational behavior. In: Pelachaud, C., Martin, J.-C., André, E., Chollet, G., Karpouzis, K., Pelé, D. (eds.) *IVA 2007*. LNCS (LNAI), vol. 4722, pp. 45–56. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74997-4_5
9. Ennis, C., O'Sullivan, C.: Perceptually plausible formations for virtual conversers. *Comput. Anim. Virtual Worlds* **23**(3–4), 321–329 (2012)
10. Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M.: Equilibrium theory revisited: mutual gaze and personal space in virtual environments. *Presence: Teleoperators Virtual Environ.* **10**(6), 583–598 (2001)
11. Zibrek, K., Kokkinara, E., McDonnell, R.: Don't stand so close to me: investigating the effect of control on the appeal of virtual humans using immersion and a proximity-based behavioral task. In: *Proceedings of the ACM Symposium on Applied Perception*, pp. 1–11. ACM, New York (2017)
12. Peters, C., Li, C., Yang, F., Avramova, V., Skantze, G.: Investigating social distances between humans, virtual humans and virtual robots in mixed reality. In: *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, pp. 2247–2249 (2018)
13. Gris, I., Novick, D.: Virtual agent interaction framework (VAIF): a tool for rapid development of social agents. In: *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 2230–2232. Springer, Berlin (2018)
14. Cafaro, A., Ravenet, B., Ochs, M., Vilhjálmsón, H.H., Pelachaud, C.: The effects of interpersonal attitude of a group of agents on user's presence and proxemics behavior. *ACM Trans. Interact. Intell. Syst. (TiiS)* **6**(2), 1–33 (2016)
15. Rodriguez, A.E., Camacho, A., Hinojos, L.J., Afravi, M., Novick, D.: A proxemics measurement tool integrated into VAIF and unity. In: *2019 International Conference on Multimodal Interaction*, pp. 508–509. Springer, Berlin (2019)
16. Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M.: Interpersonal distance in immersive virtual environments. *Pers. Soc. Psychol. Bull.* **29**(7), 819–833 (2003)
17. Iachini, T., Coello, Y., Frassinetti, F., Ruggiero, G.: Body space in social interactions: a comparison of reaching and comfort distance in immersive virtual reality. *PLoS ONE* **9**, 11 (2014)