Influence of Floor Type on Social Navigation with Small Free-Standing Groups in Virtual Reality

Sahba Zojaji[1] [0000−0003−4689−4647], Jonathan Schiött[2] [0009−0001−0445−630X], William Ivegren[2] [0009−0004−8417−6106], Andrii Matviienko[2] [0000−0002−6571−0623], and Christopher Peters[2] [0000−0002−7257−0761]

[1] The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, P.R. China
sahbazojaji@cuhk.edu.cn
[2] KTH Royal Institute of Technology, Stockholm, 11428, Sweden
{schiott,ivegren,andriim,chpeters}@kth.se

Abstract. Human footsteps play a significant role in everyday life, allowing individuals to discern the emotions, gender, and intentions of others solely from the sound of their footsteps. However, the influence of footstep sounds made when walking on different floor types in virtual reality (VR) environments when joining conversational groups remains unclear. In this paper, we present a controlled study (N=50) to assess the impact of five different floor types, associated with specific footstep sounds and visuals, on the persuasiveness of Embodied Conversational Agents (ECAs) when inviting participants to join a free-standing conversational group. We analyze routes taken by participants and the positions at which they join the group, which may be compliant or not with the agent's request when approaching the group while walking on different virtual floor types. Our findings reveal that the type of floor being walked upon, defined by footstep sounds and visual appearance, significantly impacts the persuasiveness of ECAs and the trajectories taken by participants to join the group. Participants took longer paths and joined the group in the presence of more pleasant footstep sounds. Further, they tended to adhere to social norms by avoiding walking through the group's center.

Keywords: floor type · sound · virtual reality · small free-standing groups · joining behavior.

1 Introduction

The surfaces we walk on influence not only our physical movement but also social behaviors, and this holds true in virtual environments as well, where floor types can significantly shape user experiences and interactions [20]. Although previous studies have investigated the relationship between sound design and perception in walking interactions [7, 31, 3] and have indicated that footstep sounds can convey social cues, such as emotions and intentions [3], affecting users' social behaviors and interactions [30], and cognition [7], the impact of footstep sounds on joining small free-standing groups of embodied conversational agents remains underexplored.

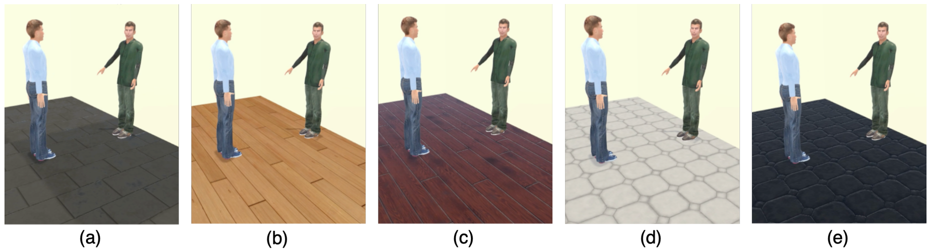

Fig. 1: We explored the impact of walking on walking on a floor that acoustically reflects: (a) no sound, (b) new wood, (c) squeaky wood, (d) cloth, and (e) metal, in virtual reality on joining a group of two Embodied Conversational Agents (ECAs).

People naturally form "free-standing conversational groups" when gathering to communicate [28]. If a newcomer wishes to join such a group politely, they should consider the social space and group [11, 13, 37]. Previous studies have explored the influence of politeness behaviors exhibited by ECAs [42] or humanoid robots [41] on the joining behaviors of individuals in small free-standing conversational groups, and their perception of the politeness of these behaviors. Another critical factor in our social navigation is the floors we are walking on and the footstep sound associated with them. People might change their social navigating behaviors on certain surfaces or spend different amounts of time walking on them. However, the impact of such footstep sounds associated with walking on different floor types in VR on individuals' joining behaviors in small groups remain unexplored. This paper investigates the impact of walking on various floor types in VR environments on individuals' social behavior. Specifically, we examine how variations in virtual floor types, associated with specific footstep sounds when walked upon and visuals, influence participants' decisions to join small groups of embodied conversational agents (ECAs) in VR (Figure 1). By manipulating the type of floor material and associated footstep sounds, namely no sound, new wood, squeaky wood, cloth, and metal, we aim to understand how the floor type influences participants' social behaviors and perceptions of social navigation in virtual environments.

In the experiment, participants faced a dilemma (Table 1) to choose between expending more effort to join a group in a socially acceptable manner or taking a shorter but socially inappropriate route. With this, we aimed to determine how likely participants are to engage in social behaviors with ECAs while walking on different floor types that are associated with different footstep sounds and visuals. Through a controlled laboratory experiment with 50 participants, we systematically examined participants' social behavior and perceptions across different floor conditions, using both objective measures and subjective assessments. Our findings revealed that floor types in VR significantly impact the persuasiveness of ECAs and the trajectories taken by participants to join the group. More specifically, participants took longer paths to join the group while walking on a floor that generated pleasant sounds (i.e., new wood). Moreover, they tended to adhere to social norms by avoiding walking through the group's center. With this work, we contribute to an empirical evaluation of how walking on various floor types in VR environments influences social behavior and navigation, with implications for designing immersive experiences that promote desired social interactions and user engagement.

2 Related Work

In this section, we outline the existing body of work about (1) free-standing groups and social proxemics and (2) the employment of footstep sounds in Virtual Reality.

2.1 Free-standing groups and social proxemics

When people come together in close physical proximity for communication or interaction, they form what is referred to as a "*free-standing conversational group*" [28]. The management of space within and among individuals in a group setting has been explored in various studies [11, 13, 15, 26, 17]. These studies extensively examine spatial dynamics within groups, with Kendon's "F-formation" theory [13] providing insights on how individuals organize themselves in such contexts. F-formation outlines the arrangement of a group where all members have equal, direct, and exclusive access to the group's social space. Within F-formation, three distinct social spaces exist: "o-space", "p-space", and "r-space". "O-space" represents a convex space enclosed by individuals engaged in social interaction, exclusive to group members. This study investigates how humans apply these spatial concepts when interacting with a group of ECAs in VR, focusing on their behavior regarding the avoidance of crossing the boundaries of "o-space" when joining the group.

Furthermore, Hall [11] categorized the area around individuals into four specific zones: (1) intimate space (0-45 cm), (2) personal space (45-120 cm), (3) social space (120-365 cm), and (4) public space (> 365 cm). According to this theory, social interactions among acquaintances mainly occur within the "social space". Therefore, in this study, we intentionally selected the onset of the social space zone [11, 12, 41] as a point presenting individuals with a dilemma: whether to navigate through the group's "o-space" or circumvent it, as they would typically do in human-human interactions. Moreover, the study seeks to determine whether participants adhere to the ECA's invitation and choose to join the particular side encouraged by the ECA.

These concepts have been implemented in computational modeling and artificial systems, particularly in scenarios where a new individual integrates into a small group. This includes computational analysis techniques [8, 27, 36], simulations of social interactions [23, 4], studies of proxemics [24, 21, 2], and datasets related to small group dynamics [40]. These endeavors have set the foundation for creating artificial models used by virtual characters and mobile robots. These agents can either act as group members accommodating newcomers or join a group while being socially aware. This awareness may include interpreting social cues or minimizing disruptions within the group [1, 33, 22, 10]. More specifically, previous studies have explored the influence of politeness behaviors exhibited by ECAs [12] or humanoid robots [41] on the joining behaviors of individuals in small free-standing conversational groups, as well as their perception of the politeness of these behaviors. Additionally, researchers have investigated the impact of multimodal communication methods on these factors [43]. However, the effect of footstep sounds associated with walking on different types of floors in VR on individuals' joining behaviors to these groups remains underexplored.

2.2 Footstep Sounds in VR

Sound plays a crucial role in human social interactions, influencing emotions, perceptions, and behaviors in diverse contexts. Footstep sounds, in particular, are integral to human perception, enabling individuals to infer body weight [30], body height [31], gender [32], social cognition [7], emotions, and intentions [3]. Several studies have investigated the intricate relationship between sound design and perception in walking interactions [39, 29, 9]. Visell et al. investigated the influence of sound design on walking interactions, providing insights into how auditory cues affect human perception and behavior [38]. Maculewicz et al. explored the impact of soundscapes on preferred walking pace, demonstrating how auditory stimuli can modulate individuals' gait and movement patterns [16]. Similarly, Turchet et al. examined how interactive sounds simulating different terrains affect walking pace, revealing the role of auditory feedback in modulating locomotion behavior [34]. Cerwén extended this research by investigating the role of soundscape design in enhancing the user experience in Japanese gardens, emphasizing the importance of considering sound in urban design [5]. Angelika et al. explored the effects of a soundscape and movement-triggered step sounds on presence in virtual reality environments, highlighting the immersive potential of audio feedback [14]. Furthermore, Tajadura-Jiménez et al. investigated the impact of altering walking sounds on perceived body weight, emotional state, and gait, revealing the profound influence of auditory feedback on embodied experiences [30]. These studies collectively underscore the multifaceted role of footstep sounds in shaping human behavior and perception in various contexts. However, in the context of virtual reality (VR) environments, understanding the impact of footstep sounds on individuals' social behavior and their perception of interactions within small free-standing conversational groups of ECAs remains a topic of ongoing exploration.

3 Methodology

We conducted a controlled laboratory experiment to systematically examine the impact of walking on various virtual floors on individuals' decisions to join a small, free-standing group of two ECAs. For this, we explored five different types of floors in VR based on participants' decisions regarding their social behaviors in response to ECA's requests to join a group at its furthest side and their perception of social navigation in the virtual environment.

3.1 Study design

To answer the research question of this study, a within-subject user study was designed with one independent variable: footstep sound. We utilized five distinct footstep sounds associated with walking on five distinct types of virtual floors in VR: *no sound*, *new wood*, *squeaky wood*, *cloth*, and *metal* (Figure 1). This selection aimed to explore the impact of walking on various types of floors in VR and their associated visuals on human social behavior while being invited to join a free-standing, conversational small group of ECAs. Additionally, it aimed to examine potential differences between natural (no sound, new wood, and squeaky wood) and unnatural footstep sounds (cloth and metal). Existing literature [35, 14] suggests that footstep sounds influence VR on user behavior, prompting our consideration of this comprehensive evaluation.

Our pilot tests showed that the most accurate synchronization of footstep sounds occurred when an experimenter manually played sounds corresponding to participants' footsteps. The trained experimenter monitored walking patterns in VR and triggered sounds to match each step, which would also increment the step count for each trial. During the experiment, participants walked barefoot to minimize the sound of their footsteps from the real world.

The experiment involved positioning two ECAs in a face-to-face group formation, with a distance of 125 cm separating them, within a room measuring 4.5 by 7 meters (Figure 2). This distance marks the initial boundary within social space [11] and is based on prior research [12] as a point that presents individuals with a dilemma: whether to navigate through the group's *o-space* or circumvent it. The main ECA (A1) was positioned facing the participants and extended an invitation to join the group by employing a combination of verbal and non-verbal behavior. Throughout the entire duration of each trial, A1 consistently maintained eye contact with the participants. Moreover, the secondary ECA (A2) also initiated and sustained eye contact as participants approached the group. The design of this approach was informed by prior studies [12], aiming to establish a welcoming and inclusive environment for participants to join the group and to demonstrate the ECAs' welcoming attitude, expectation, and preparedness for participants to join their group.

To systematically explore the impact of footstep sound associated with walking on different floor types in VR on participants' social behavior when joining groups of ECAs, a within-subject user experiment consisting of 5 trials in VR was designed. We counterbalanced the order of the floors using a Balanced Latin square. To mitigate the potential influence of gender on participant responses, female agents exclusively interacted with female participants, while male agents engaged with male participants based on previous literature [41]. Given that there is a significant difference in interaction between the agents' gender and the participants' gender, we aimed to exclude this bias from this experiment [18]. Finally, adopting a within-subjects design enabled participants to experience and compare all conditions, reducing individual variability and enhancing statistical robustness [6].

3.2 Social dilemma

There were a number of options that participants had in relation to whether they would join the group and how they did so. If they decided to join the group, they needed to choose both a position at which they would join as well as the route to get there. This resulted in four possibilities (as outlined in Table 1):

1. Opting for a socially acceptable but more effortful approach to join the group, complying with the ECA's request. This involved taking an *inconvenient* route around the group, which required approximately 12 steps.

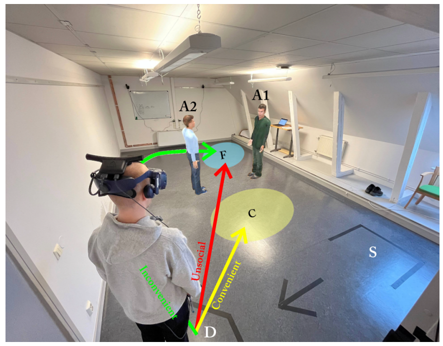

Fig. 2: **Experiment room:** *S* represents the starting spot for each trial where participants started their movements, and *D* represents the spot where participants should decide whether they want to continue walking on a specific type of floor and join the group or not. The *Closest* (*C*) and *Furthest* (*F*) sides of the group for joining from *D* are distinguished by yellow and blue circles, respectively. Additionally, three distinct hypothetical routes for joining the group are depicted, each associated with a specific color: green for *inconvenient*, red for *unsocial*, and yellow for *convenient*. In all trials, A1 (main ECA) invited participants to join at the *furthest* side (*F*). A participant is standing at the *D* spot and fully immersed in VR.

2. Choosing an *unsocial* route to comply with the ECA's request, by walking directly through the center of the group, violating the group's *o-space*. This option required about 8 steps but still adhered to the ECA's request.
3. Selecting a *convenient* route, involving only about 4 steps to join the group at the *closest* side (Figure 2). This choice balanced effort and social acceptance but conflicted with the ECA's invitation to join at the *furthest* side.
4. Selecting the *effortless* approach by not walking to join the group and ignoring the ECA's request completely. This required no movement toward the group or 0 steps to approach the ECAs.

This setup allowed for examining participants' priorities and joining behavior, considering their decision to join in response to persuasion and complying with the ECA's request, and their approach to doing so while adhering to social norms while walking, which involved varying levels of effort. The social acceptability of these potential routes was based on established social norms and derived from Kendon's theory [13]. Kendon's theory suggests that people generally avoid walking through a group's "o-space" when adequate space exists around it. The specific dilemmas presented aimed to investigate how individuals manage conflicting social norms while considering invitations from ECAs and the associated effort implications. Table 1 briefly summarizes the four dilemmas, encapsulating the varying trade-offs related to social norms, levels of persuasion in politeness behaviors, and effort considerations.

Furthermore, the dilemma indicates how individuals perceive ECAs as social entities akin to humans and conform to their requests. It also explores whether the invitation by the ECAs could impact participants' decisions regarding their behavior in walking

Table 1: Various alternatives and their associated trade-offs from the participants' perspective, as illustrated in Figure 2.

Routes	Persuasion	Social adherence	Effort
Effortless	No	No	Nothing
Convenient	No	Yes	Low
Unsocial	Yes	No	Medium
Inconvenient	Yes	Yes	High

on different types of floors and joining the group. Moreover, participants were not incentivized to adhere to the requests made by the ECA to abstain from walking through the group's *o-space* or even to join the group. In turn, they had complete freedom to decide whether they wanted to walk on each specific floor and join the group, where and how they wished to join the group, and none of these potential routes were presented to them. Even when standing on the D spot, they were reminded in the virtual environment that they had the choice not to walk anymore on that specific type of floor and end the trial or walk and join the group.

3.3 Apparatus

Using the Unity 3D game engine[3], we constructed a virtual indoor room that faithfully replicated the features of the physical room where the study took place. The virtual starting spot, S, and the decision spot, D, were accurately aligned with the physical environment to maintain a consistent spatial reference. In the virtual indoor room of the experiment, two virtual characters, specifically the Greta virtual agent[4] [19], were positioned at the room's center. Speech generation in these virtual characters was accomplished using CereProc text-to-speech[5]. Two male and two female Greta agents were deployed to account for potential gender effects. Female participants interacted exclusively with female characters, while male participants interacted exclusively with male characters. Additionally, trajectory data from the VR headset was recorded throughout the study. To enable participants' unrestricted mobility, we equipped the headsets with HTC VIVE wireless adapters powered by wearable battery packs (Figure 2). Participants wore the VR headset throughout the trials and could end each trial by verbally saying “End”.

3.4 Measures

Participants' *joining behavior* is assessed through six objective measures:

1. **Persuasiveness:** The study captured participants' compliance with the ECA's invitations to join the group at the *furthest* side in each trial.

[3] https://www.unity.com/
[4] https://github.com/isir/greta/wiki
[5] TTS: https://www.cereproc.com/en/home

2. **Social Adherence:** Participants opting to join at the *furthest* side faced a choice: either traversing between the two ECAs and violating the group's "o-space," or circumventing them while respecting this social boundary. Analysis of this data quantifies participants' conformity to social norms in their joining behavior.
3. **Number of Approach Steps:** This metric quantifies the number of steps taken by participants from their decision spot (depicted as "D" in Figure 2) to the point where they terminated the trial by stating "end," signifying their perception of joining the group.
4. **Approach Time:** This parameter indicates the duration (measured in seconds) of each trajectory, spanning from the commencement of each trial at the decision spot ("D") to the point where participants terminated the trial.
5. **Distance Traveled:** This metric encompasses the length of participants' trajectories (measured in meters) from their decision spot ("D") to the point where they terminated the trial.
6. **Distance to Group:** This measure records the distance (in meters) between the point where participants halted their movement and the center of the group.

In addition to these objective metrics, we also administered a questionnaire at the conclusion of each trial for five perceptual measures related to participants' perception in VR:

1. **Politeness:** "The agent's request to invite me to join the group on this type of floor was polite."
2. **Disturbance:** "It felt like I was disturbing the agents while walking on this type of floor."
3. **Effort:** "The effort I experienced while walking on this type of floor was high."
4. **Naturalness:** "The sound of my footsteps on this type of floor felt natural."
5. **Awareness:** "I was aware of the sound of my footsteps on this type of floor."

Each question was tailored to evaluate one of the five dependent variables concerning the participant's perception of their footstep sounds in VR during the task. These variables include the *politeness* of the ECA's request to invite participants to join the group at its furthest side on different types of floors, the *disturbance* caused by participants for the ECAs while walking on different types of floors, the level of *effort* they exerted, the *naturalness* of the footstep sound, and the participants' awareness of their footstep sounds. Participants were asked to indicate their level of agreement with the provided statements on a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree". Furthermore, participants were requested to assess their characteristics regarding *rules obedience* and *social role dynamics* after the study. They indicated their level of agreement with the following statements on a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree".

1. "I follow rules even when faced with unpleasant consequences for myself."
2. "I follow rules even when others encounter undesirable outcomes."
3. "I see myself as someone who is influential while I am with a group of people."
4. "I see myself as someone who is a follower while I am with a group of people."

3.5 Procedure

At the beginning of the experiment, participants were instructed to remove their shoes to eliminate any potential interference from their real footstep sounds while walking in the physical room with their virtual footstep sounds in the virtual room. Following this, participants' demographic data and signed consent forms were collected. Subsequently, participants wore the VR headset and became fully immersed in the virtual environment. We then introduced them to the experimental setup, which included the virtual room and ECAs. Participants were informed that all the ECAs were entirely autonomous. To familiarize themselves with the environment, participants completed five demo trials associated with five different types of floors.

For each trial of the experiment, participants were fully immersed in VR and instructed to move to the designated starting spot (S) at the outset (Figure 2). From there, they were required to walk towards the decision spot (D), covering approximately five steps. This phase aimed to familiarize them with the footstep sounds associated with walking on a specific type of floor for each trial. Upon reaching the decision spot, participants encountered a reminder embedded in the VR environment, prompting them of their two choices: 1) to continue walking on the current floor and join the group, or 2) to cease walking on that particular floor and conclude the trial. Subsequently, the trial commenced, and the main ECA, A1, invited participants to join the group at its furthest side by verbally stating "Come here!" while pointing towards the specific side where the ECA wanted the participants to join (see Figure 1). Participants' position and orientation at the decision spot (D) were carefully designed so that they faced the group of ECAs directly in front of them. This positioning ensured participants had a clear view of the ECAs and their invitations, allowing them to easily perceive and respond to the agent's requests to join the group.

A beep sound played within the VR environment, initiated by the experimenter from the control room, signaled the start of each trial before participants began to move towards the group or conclude the trial. This sound ensured that the main ECA (A1) had finished its invitation, allowing participants to act accordingly. Participants were instructed to conclude each trial by verbally stating "End," either at the decision spot or after joining the group. After concluding each trial, participants were instructed to return to their starting spot, "S". They were required to respond to the five perceived footstep sound questions within the VR environment. Once they had completed this task, they were already at the initial location and could proceed to start the next trial. There were no specific requirements regarding how they should join the group, and they were free to do so as they preferred. Additionally, participants had the complete freedom to end the trial at the decision point and choose not to join the group. The initial positioning and angles of both the group and the participants from the decision point were chosen to facilitate a direct, least effort, and *convenient* path, enabling participants to join the group at the *closest* side with minimal effort, should they choose to do so. However, A1 consistently invited participants to the opposite side of the group, requiring a more effortful and *inconvenient* route to reach, approximately twice the distance.

After the experiment, participants were asked to complete a post-study survey to offer feedback on their overall experience and characteristics regarding *rules obedience* and *social role dynamics*. Additionally, they were thanked for their participation and received gift cards or cinema tickets as tokens of appreciation. Furthermore, they were debriefed about the purpose of the study.

3.6 Participants

A total of 50 participants (29 male, 21 female), ranging in age from 20 to 66 years ($M = 29, SD = 11$), and proficient in English were recruited for this study. Among them, 94% reported having little or no prior experience with VR, and 94% had infrequent VR usage, typically a few times a year. Furthermore, all participants had good hearing without any significant hearing impairment. The data collected from participants were anonymized, and informed written consent was obtained from each participant before the experiment began. Each participant completed 5 trials, resulting in a total of 250 trials used for the final analysis.

3.7 Data analysis

Given the non-parametric nature of the collected data, we applied the Friedman test as an omnibus test and Conover test with Bonferroni correction for the post-hoc analysis.

4 Results

In this Section, results are presented for the measures described in Section 3.4: six objective measures related to participants' joining behavior and five perceptual measures.

4.1 Objective measures

Persuasiveness and social adherence The Figure 3 illustrates how the floor type influenced the number of times participants joined the group of agents at the furthest side while walking between the two agents (by taking an *unsocial* route) or walking around them (by taking an *inconvenient* route). Additionally, they had an option to join at the closest side (by taking a *convenient* route) and do not join at all. We illustrate these behaviors in Figure 3.

Number of approach steps The number of approach steps participants took was the highest with the noiseless floor ($Md = 11, IQR = 3$), followed by the new wood ($Md = 10.5, IQR = 3$), cloth ($Md = 10, IQR = 4$), squeaky wood ($Md = 7.5, IQR = 11$), and metal ($Md = 0, IQR = 8$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 80.9, p < 0.001, \eta^2 = 0.4$). The post-hoc analysis has shown statistically significant differences between all pairs, except for new wood and cloth ($p = 0.55$): squeaky wood and metal ($p = 0.0016$) and the remaining pairs ($p < 0.001$) (Figure 4).

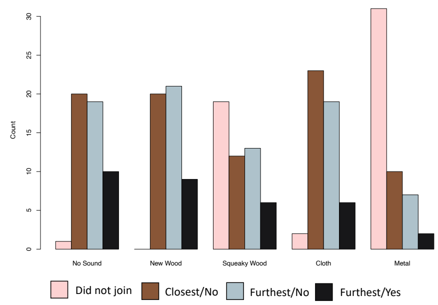

Fig. 3: The count of instances where participants joined the group at the closest side without crossing the o-space via a *convenient* route (*Closest/No*), joined at the furthest side without crossing the o-space via an *inconvenient* route (*Furthest/No*), joined at the furthest side while crossing the o-space of the group via an *unsocial* route (*Furthest/Yes*), and did not join the group at all.

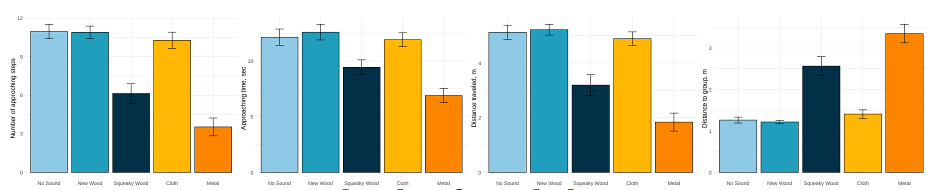

Fig. 4: Objective measures of participants' social behaviors for their joining behaviors in VR.

Approach time The amount of time participants took to approach the group was the highest with the new wood ($Md = 11.56sec, IQR = 3.48$), followed by cloth ($Md = 11.36sec, IQR = 3.4$), noiseless floor ($Md = 11.2sec, IQR = 2.99$), squeaky wood ($Md = 10.3sec, IQR = 7.78$), and metal ($Md = 4.45sec, IQR = 6.5$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 35, p < 0.001, \eta^2 = 0.175$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$), except for noiseless floor and new wood ($p = 0.99$), noiseless floor and cloth ($p = 0.99$) and new wood and cloth ($p = 0.99$) (Figure 4).

Distanced traveled The total distance traveled was the highest with the new wood ($Md = 5.39m, IQR = 2.19$), followed by noiseless floor ($Md = 5.29m, IQR = 2.19$), cloth ($Md = 5.08m, IQR = 2.24$), squeaky wood ($Md = 3.7m, IQR = 5.49$), and metal ($Md = 0.2m, IQR = 3.64$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 55.9, p < 0.001, \eta^2 = 0.28$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$), except for noiseless floor and new wood ($p = 0.074$), noiseless floor and cloth ($p = 0.44$) and new wood and cloth ($p = 0.99$) (Figure 4).

Distance to group The distance to group was the highest with the metal ($Md = 4.49m, IQR = 3.2$), followed by new wood ($Md = 1.19m, IQR = 0.26$), noiseless floor ($Md = 1.19m, IQR = 0.16$), squeaky wood ($Md = 1.41m, IQR = 3.3$), and cloth ($Md = 1.22m, IQR = 0.18$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 48, p < 0.001, \eta^2 = 0.24$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$), except for noiseless floor and new wood ($p = 0.99$), noiseless floor and cloth ($p = 0.79$) and new wood and cloth ($p = 0.28$) (Figure 4).

4.2 Perceptual measures

Politeness of the request The perceived politeness was the highest with the noiseless floor ($Md = 3, IQR = 2$), followed by new wood ($Md = 3, IQR = 2$), cloth ($Md = 3, IQR = 2$), squeaky wood ($Md = 3, IQR = 2$), and metal ($Md = 3, IQR = 1$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 14.5, p = 0.005, \eta^2 = 0.07$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$), except for noiseless floor and new wood ($p = 0.99$), noiseless floor and cloth ($p = 0.32$) and new wood and cloth ($p = 0.99$) (Figure 5).

Disturbing the ECA The perceived disturbance of agents was the highest with the metal floor ($Md = 5, IQR = 1$), followed by squeaky wood ($Md = 4, IQR = 2$), cloth ($Md = 2, IQR = 1$), new wood ($Md = 2, IQR = 1$), and noiseless floor ($Md = 1, IQR = 0$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 132.3, p < 0.001, \eta^2 = 0.66$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$), except for new wood and cloth ($p = 0.99$) (Figure 5).

Effort The perceived effort was the highest with the metal floor ($Md = 4, IQR = 1$), followed by squeaky wood ($Md = 4, IQR = 1$), cloth ($Md = 2, IQR = 1$), new wood ($Md = 1, IQR = 1$), and noiseless floor ($Md = 1, IQR = 0$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 133.1, p < 0.001, \eta^2 = 0.66$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$) (Figure 5).

Naturalness of the sound The floor sound was perceived as the most natural with the noiseless floor ($Md = 4, IQR = 1$), followed by new wood ($Md = 4, IQR = 1.5$), cloth ($Md = 2, IQR = 1$), squeaky wood ($Md = 2, IQR = 1.75$), and metal ($Md = 1, IQR = 0$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 116.2, p < 0.001, \eta^2 = 0.58$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$) (Figure 5).

Awareness of the footsteps The awareness of the footsteps was the highest with the metal floor ($Md = 5, IQR = 0$), followed by squeaky wood ($Md = 5, IQR = 0$), cloth ($Md = 4, IQR = 1$), new wood ($Md = 4, IQR = 1$), and noiseless floor ($Md = 2, IQR = 2$). This finding was supported by the statistically significant main effect for the type of floors ($chi^2(4) = 127.5, p < 0.001, \eta^2 = 0.63$). The post-hoc analysis has shown statistically significant differences between all pairs ($p < 0.001$), except for new wood and cloth ($p = 0.99$) and squeaky wood and metal ($p = 0.33$) (Figure 5).

4.3 Human factors

Note that we examined the collected data for indicators of rule obedience and social role dynamics. However, our analysis did not reveal any significant differences among participants.

5 Discussion

5.1 Pleasant and Natural Sounds Lead to Polite Behaviors

The study's results provided insights into the intricate relationship between auditory cues, social behavior, and navigation in immersive virtual reality environments. The findings indicate that participants exhibit nuanced social behaviors influenced by the type of footstep sounds associated with different types of floors. Specifically, participants preferred joining a small group in the presence of pleasant footstep sounds, such as those associated with new wood floors. Participants tended to join the group at its furthest side when the footstep sounds were pleasant, such as with new wood floors. Conversely, they joined more to the closest side when encountering noisy or unpleasant but natural footstep sounds, such as with squeaky wood floors. In situations where the footstep sound was not natural, such as with metal, they mainly preferred not to join the group at all. This observation suggests that participants' joining behavior is influenced by the perceived pleasantness of footstep sounds and their perceived naturalness. This also extends the existing body of work about joining small groups that has indicated that politeness plays more important role when influencing people's behavior rather than visual representation of virtual characters [41].

When encountering pleasant footstep sounds, participants may feel more comfortable engaging with the group and, therefore, choose to join at the furthest side and take longer routes. Conversely, when faced with noisy or unnatural footstep sounds, participants may prioritize minimizing discomfort, leading them to opt for the closest side or avoid joining the group altogether. This suggests that auditory cues play a significant role in shaping social navigation decisions, with individuals being more inclined to engage in social interactions when presented with pleasant auditory stimuli. Furthermore, the study highlights the importance of considering social norms in virtual environments. Participants tended to adhere to social norms by avoiding walking through the group's center, even without explicit instructions. This underscores the role of social cues and conventions in guiding behavior, even in virtual contexts where social interactions may differ from real-world scenarios.

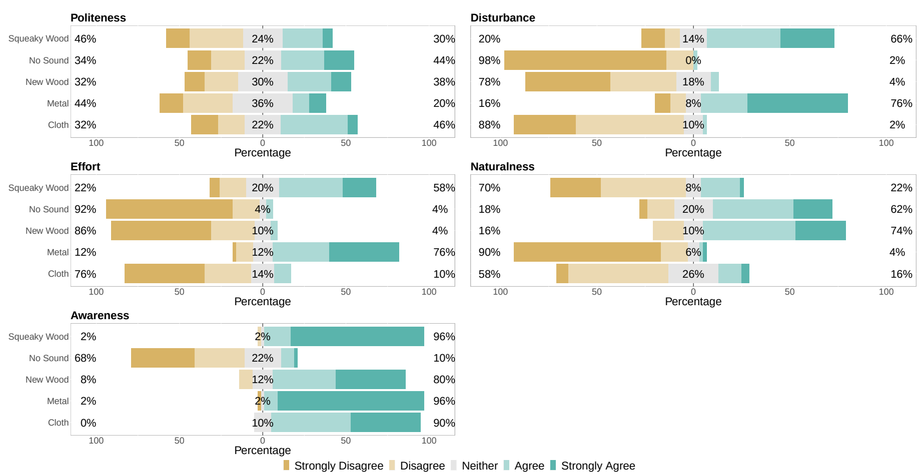

Fig. 5: Perceived politeness of the request, disturbing the ECA, effort, Naturalness of the sound, Awareness of the footsteps. Strongly disagree indicates low politeness/disturbance/effort/naturalness/awareness, and strongly agree – high politeness/disturbance/effort/naturalness/awareness.

5.2 Floor Visuals vs. Floor Noises

In this experiment, we explored the augmentation of floor visuals with noises created by these floors. Since participants' behavior and decision-making were not influenced by pure visuals of the floors they were walking on, we discovered that the floor's noises played a decisive role. Since we focused on the indoor interactions between agents, we explored wooden floors as one of the common examples of indoor flooring as a natural and pleasant option. Consequently, unpleasant and unnatural sounds lead to shorter walking distances to decrease annoyance. With this work, we essentially aimed to explore the sounds from the opposite sides of the spectrum of annoyance to create a basis for exploring the acoustic and visual mismatch to prime participants in virtual environments when necessary. For example, this knowledge helps future designers create VR experiences and behaviors that demonstrate the importance of comfort during interaction rather than adherence to social norms. The sole presence of the footsteps sounds in the absence of visual feedback would lead to similar behaviors and reactions, e.g., unpleasant sounds leading to shorter walking distances, which needs to be further explored in future studies. Similarly to the previous work [16, **?**] that explored walking paces and patterns depending on the footstep sounds.

5.3 Design implications

The findings of this study are relevant especially to research concerning social interactions with virtual characters in virtual reality. While such research typically focuses on the appearance and animation quality of the characters, our study suggests that attention also should be paid to the design of the environment in which the interaction takes place. For example, cost effective full-body virtual character motion solutions that require limited sensing [25] are becoming increasingly viable for representing self-avatars in virtual environments. It is natural to add footstep sounds to enhance the feeling of synchronised motions when walking in order to enhance the virtual experience. Our study suggests that these sounds should be carefully chosen as they may have unintended impacts on how participants engage with and relate to social interactions with virtual characters in those environments. Beyond avoiding unintended consequences of poor sound choices, it might also be possible to use sounds in order to purposefully flavour social interactions in virtual environments. Additionally, altering the starting positions of participants and exploring various group formations of ECAs could provide further insights into social behavior and interaction patterns in immersive virtual environments.

6 Conclusion

This study investigated the influence of auditory cues on social behavior and navigation in immersive virtual reality environments. Our results indicate that participants demonstrated a preference for joining small groups in the presence of pleasant footstep sounds, highlighting the importance of footstep sounds in shaping social navigation decisions in VR. Future research could explore additional factors influencing social navigation in virtual environments, such as visual cues, interpersonal dynamics, various starting positions, and individual differences in auditory perception. Overall, this study contributes to our understanding of how auditory cues shape social behavior and navigation in virtual reality, highlighting the importance of considering multisensory factors in the design of immersive environments in a social VR context.

References

1. Althaus, P., Ishiguro, H., Kanda, T., Miyashita, T., Christensen, H.I.: Navigation for human-robot interaction tasks. In: IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004. vol. 2, pp. 1894–1900. IEEE (2004)
2. Blascovich, J., Loomis, J., Beall, A.C., Swinth, K.R., Hoyt, C.L., Bailenson, J.N.: Immersive virtual environment technology as a methodological tool for social psychology. Psychological inquiry **13**(2), 103–124 (2002)
3. Bresin, R., de Witt, A., Papetti, S., Civolani, M., Fontana, F., et al.: Expressive sonification of footstep sounds. Proceedings of ISon **2010**, 3rd (2010)
4. Cafaro, A., Ravenet, B., Ochs, M., Vilhjálmsson, H.H., Pelachaud, C.: The effects of interpersonal attitude of a group of agents on user's presence and proxemics behavior. ACM Transactions on Interactive Intelligent Systems (TiiS) **6**(2), 1–33 (2016)
5. Cerwén, G.: Listening to japanese gardens ii: expanding the soundscape action design tool. Journal of urban design **25**(5), 607–628 (2020)
6. Charness, G., Gneezy, U., Kuhn, M.A.: Experimental methods: Between-subject and within-subject design. Journal of economic behavior & organization **81**(1), 1–8 (2012)
7. Clausen, S., Tajadura-Jiménez, A., Janssen, C.P., Bianchi-Berthouze, N.: Action sounds informing own body perception influence gender identity and social cognition. Frontiers in Human Neuroscience **15**, 688170 (2021)
8. Cristani, M., Bazzani, L., Paggetti, G., Fossati, A., Tosato, D., Del Bue, A., Menegaz, G., Murino, V.: Social interaction discovery by statistical analysis of f-formations. In: BMVC. vol. 2, p. 4 (2011)
9. Dietz, F., Hoppe, M., Karolus, J., Wozniak, P.W., Schmidt, A., Machulla, T.: Vrsneaky: Stepping into an audible virtual world with gait-aware auditory feedback. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–4. CHI EA '20, Association for Computing Machinery, New York, NY, USA (2020)
10. Gao, Y., Yang, F., Frisk, M., Hernandez, D., Peters, C., Castellano, G.: Social behavior learning with realistic reward shaping. arXiv preprint arXiv:1810.06979 (2018)
11. Hall, E.T.: The hidden dimension, vol. 609. Garden City, NY: Doubleday (1966)
12. Iop, A., Zojaji, S., Peters, C.: Don't walk between us: adherence to social conventions when joining a small conversational group of agents. In: Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents. pp. 1–8 (2022)
13. Kendon, A.: Conducting interaction: Patterns of behavior in focused encounters, vol. 7. Cambridge University Press (1990)
14. Kern, A.C., Ellermeier, W.: Audio in vr: Effects of a soundscape and movement-triggered step sounds on presence. Frontiers in Robotics and AI **7**, 20 (2020)
15. Li, R., van Almkerk, M., van Waveren, S., Carter, E., Leite, I.: Comparing human-robot proxemics between virtual reality and the real world. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 431–439 (2019)
16. Maculewicz, J., Erkut, C., Serafin, S.: How can soundscapes affect the preferred walking pace? Applied Acoustics **114**, 230–239 (2016)
17. Mumm, J., Mutlu, B.: Human-robot proxemics: Physical and psychological distancing in human-robot interaction. In: Proceedings of the 6th International Conference on Human-Robot Interaction. p. 331–338. HRI '11, Association for Computing Machinery, New York, NY, USA (2011)
18. Nag, P., Yalçın, O.N.: Gender stereotypes in virtual agents. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents. IVA '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3383652.3423876, https://doi.org/10.1145/3383652.3423876

19. Niewiadomski, R., Bevacqua, E., Mancini, M., Pelachaud, C.: Greta: an interactive expressive eca system. In: AAMAS '09: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems. vol. 2, p. 1399–1400. International Foundation for Autonomous Agents and Multiagent Systems (May 2009), https://dl.acm.org/doi/abs/10.5555/1558109.1558314
20. Nilsson, N.C., Serafin, S., Steinicke, F., Nordahl, R.: Natural walking in virtual reality: A review. Computers in Entertainment **16**(2) (Apr 2018)
21. Novick, D., Rodriguez, A.E.: A Comparative Study of Conversational Proxemics for Virtual Agents, Lecture Notes in Computer Science, vol. 12770, p. 96–105. Springer International Publishing, Cham (2021)
22. Pathi, S.K.: Join the group formations using social cues in social robots. In: Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems. pp. 1766–1767. International Foundation for Autonomous Agents and Multiagent Systems (2018)
23. Pedica, C., Vilhjálmsson, H.H.: Study of nine people in a hallway: Some simulation challenges. In: the 18th Int. Conf. on Intelligent Virtual Agents. pp. 185–190. ACM (2018)
24. 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. p. 2247–2249. AAMAS '18, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2018)
25. Ponton, J.L., Yun, H., Aristidou, A., Andujar, C., Pelechano, N.: Sparseposer: Real-time full-body motion reconstruction from sparse data. ACM Trans. Graph. **43**(1) (oct 2023). https://doi.org/10.1145/3625264, https://doi.org/10.1145/3625264
26. Samarakoon, S.B.P., Muthugala, M.V.J., Jayasekara, A.B.P.: A review on human–robot proxemics. Electronics **11**(16), 2490 (2022)
27. Setti, F., Lanz, O., Ferrario, R., Murino, V., Cristani, M.: Multi-scale f-formation discovery for group detection. In: 2013 IEEE International Conference on Image Processing. pp. 3547–3551. IEEE (2013)
28. Setti, F., Russell, C., Bassetti, C., Cristani, M.: F-formation detection: Individuating free-standing conversational groups in images. PloS one **10**(5) (2015)
29. Son, H., Gil, H., Byeon, S., Kim, S.Y., Kim, J.R.: Realwalk: Feeling ground surfaces while walking in virtual reality. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. p. 1–4. CHI EA '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3170427.3186474, https://doi.org/10.1145/3170427.3186474
30. Tajadura-Jiménez, A., Basia, M., Deroy, O., Fairhurst, M., Marquardt, N., Bianchi-Berthouze, N.: As light as your footsteps: altering walking sounds to change perceived body weight, emotional state and gait. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. pp. 2943–2952 (2015)
31. Tajadura-Jiménez, A., Deroy, O., Marquardt, T., Bianchi-Berthouze, N., Asai, T., Kimura, T., Kitagawa, N.: Audio-tactile cues from an object's fall change estimates of one's body height. Plos one **13**(6), e0199354 (2018)
32. Tajadura-Jiménez, A., Newbold, J., Zhang, L., Rick, P., Bianchi-Berthouze, N.: As light as you aspire to be: Changing body perception with sound to support physical activity. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–14 (2019)
33. Truong, X.T., Ngo, T.D.: "to approach humans?": A unified framework for approaching pose prediction and socially aware robot navigation. IEEE Transactions on Cognitive and Developmental Systems **10**(3), 557–572 (2018)

34. Turchet, L., Serafin, S., Cesari, P.: Walking pace affected by interactive sounds simulating stepping on different terrains. ACM Transactions on Applied Perception (TAP) **10**(4), 1–14 (2013)
35. Umair Bin Altaf, M., Butko, T., Juang, B.H.: Acoustic gaits: Gait analysis with footstep sounds. IEEE Transactions on Biomedical Engineering **62**(8), 2001–2011 (2015). https://doi.org/10.1109/TBME.2015.2410142
36. Vázquez, M., Steinfeld, A., Hudson, S.E.: Parallel detection of conversational groups of free-standing people and tracking of their lower-body orientation. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 3010–3017. IEEE (2015)
37. Vilhjálmsson, H.H.: Interaction in social space. In: The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application, pp. 3–44 (2022)
38. Visell, Y., Fontana, F., Giordano, B.L., Nordahl, R., Serafin, S., Bresin, R.: Sound design and perception in walking interactions. International Journal of Human-Computer Studies **67**(11), 947–959 (2009)
39. Visell, Y., Law, A., Ip, J., Smith, S., Cooperstock, J.R.: Interaction capture in immersive virtual environments via an intelligent floor surface. In: 2010 IEEE Virtual Reality Conference (VR). pp. 313–314. IEEE (2010)
40. Yang, F., Gao, Y., Ma, R., Zojaji, S., Castellano, G., Peters, C.: A dataset of human and robot approach behaviors into small free-standing conversational groups. PloS one **16**(2), e0247364 (2021)
41. Zojaji, S., Matviienko, A., Leite, I., Peters, C.: Join me here if you will: Investigating embodiment and politeness behaviors when joining small groups of humans, robots, and virtual characters. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. CHI '24, Association for Computing Machinery, New York, NY, USA (2024). https://doi.org/10.1145/3613904.3642905, https://doi.org/10.1145/3613904.3642905
42. Zojaji, S., Peters, C., Pelachaud, C.: 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. p. 1–8. ACM (Oct 2020). https://doi.org/10.1145/3383652.3423917
43. Zojaji, S., Červeň, A., Peters, C.: Impact of multimodal communication on persuasiveness and perceived politeness of virtual agents in small groups. In: Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents. pp. 1–8 (2023)