IADIS International Journal on Computer Science and Information Systems
Vol. 18, No. 2, pp. 194-212
ISSN: 1646-3692

AMPLIFICATION OF CO-PRESENCE IN GROUP VISITS TO VIRTUAL HOUSES WITH GRAPHICAL REPRESENTATIONS FROM VIDEO GAMES

André Freitas[1], Filipe Santos[1,2] and Pedro Santana[1,2]
[1]*ISCTE – Instituto Universitário de Lisboa, Portugal*
[2]*ISTAR – Information Sciences, Technologies and Architecture Research Center, Portugal*

ABSTRACT

An individual can observe one's surrounding space. The space can contain various elements (e.g., objects, furniture), which in turn convey visual information to the individual, giving the sense of being present within the space. The space can be filled by additional individuals who also aim to observe their surroundings. The human ability to observe and interact with the surrounding elements plays a vital role in creating a feeling of being socially present and connected in a physical space, also know as co-presence. In the virtual world, similar sensations of co-presence are often conveyed through graphical representations. These representations, along with other methods aimed at enhancing co-presence (such as interfaces designed to foster a sense of shared presence), have been extensively researched and utilized in single player and multiplayer video games. The present study explores the application of some of those strategies to enhance the feeling of co-presence among individuals during virtual group house tours, which is a relevant application scenario to the real estate industry and architectural participatory design. To this aim, a tool was developed and evaluated in a user study involving 33 participants. The findings indicate that all the implemented strategies effectively enhance the sense of co-presence within the virtual environment and encourage meaningful collaborative interactions.

KEYWORDS

Co-Presence, Virtual Environment, Videogames, Collaborative Interaction

1. INTRODUCTION

In an ideal scenario, users engaging with virtual environments should experience complete presence and immersion. However, these virtual experiences come with constraints, being unable to entirely replicate the advantages of real-life encounters. Nowadays, technical concerns persist, such as effectively addressing issues like cyber-sickness, raising the need for continued research in this field.

Effective spatial navigation support is crucial within virtual environments as unassisted navigation can prove challenging and potentially causing user disorientation (Vinson, 1999), especially in Collaborative Virtual Environments (CVE) (Casanueva and Blake, 2000), as these allow users to share the same virtual space and perform collaborative tasks. In these environments, users need to have a sense of co-presence, otherwise they will not be able to collaborate. Digital avatars of the users, i.e., graphical representations of the users, are often employed for this purpose (Churchill and Snowdon, 1998) (Peña Pérez Negrón et al., 2020).

Extensive research has been conducted on digital avatars, aiming at enhancing co-presence, such as the use of graphical representations that induce a sense of shared presence. These methods are found widespread in multiplayer videogames (Warpefelt, 2016) (Steuer, 1992). Over the years, the videogame industry has expanded the realm of possibilities in crafting different approaches that entice users to dedicate their time to explore the extensive offerings of virtual environments, encompassing recreational and educational aspects or both. Videogames are a special case of CVE, many of them capable of providing the players with a powerful sense of co-presence when facing challenging complex collaborative navigation tasks. The value of studying videogames beyond their entertainment role is well established. For instance, previous studies addressed the role of videogames as learning tools in specific age groups (Fernandes et al., 2023) and their effects on physical and mental health of the users (Stanney and Cohn, 2009).

Real estate agencies are swiftly embracing the technological age by introducing innovative and user-friendly methods for property viewing. Nowadays, individuals can virtually explore the interiors of houses through web browsers, facilitated by hardware capable of capturing 3D images of the houses themselves. This technology creates the sensation of visiting a house inside one's computer. However, there are limitations to user navigation, restricted to specific viewpoints where the 3D images were captured. Virtual environments hold immense potential as a novel approach to visit existing houses or future projects, offering unrestricted navigation and, most importantly, a wide array of functions that can transform it into a user-centric ecosystem, ensuring ease of use and accessibility.

In this paper, we study the hypothesis that some of the co-presence strategies employed by videogames are capable of amplifying co-presence in virtual group house visits. Engaging in virtual group house tours holds significant importance within the real estate sector. The capability to preview homes virtually not only saves time and costs but also mirrors the collective process people often undertake when selecting a house. Conducting group visits to virtual properties also serves as a valuable tool in participatory architectural design. It allows clients to collectively explore initial house designs, enabling them to offer timely feedback that informs and shapes subsequent design iterations.

We have designed and implemented a Unity-based CVE, hereafter VVApp (from Virtual Visit App), in which remote users are able to jointly visit virtual houses, with co-presence amplified by means of graphical representations borrowed from videogames. A user-study with 33 participants was conducted, to validate our hypothesis. During a series of tasks, user behavior was monitored to evaluate the effectiveness of the implemented co-presence strategies in facilitating collaborative navigation, communication, and decision-making. The obtained results show that the implemented strategies enrich the sense of co-presence in the virtual environment and trigger interesting collaborative interactions.

The structure of this paper is as follows: Section 2 discusses related work, followed by a description of the developed CVE in Section 3. Section 4 covers the experimental setup, while Section 5 presents its results. The conclusions and suggestions for future research are outlined in Section 6.

2. RELATED WORK

Casanueva and Blake (2000) studied the effects that avatars have on co-presence in a CVE. The authors implemented two user studies, being the appearance experience and the function experience. The conclusion reached by the authors was based on the presence and co-presence scores and survey responses, which turned out to be quite positive. Buck et al. (2019) observed that co-presence can exist in a CVE at a level remarkably similar to that of reality. The authors led a user study that analyzes the common actions of two users crossing a fake corridor in real life and in virtual reality, considering the gender of the user and avatar. A wide variety of body adjustments and single rows for each gender were registered in this experiment. Podkosova and Kaufmann (2018) concluded that users try not to collide in the virtual world even when aware that t hey were safe from real collisions. Pimentel and Vinkers (2021) refer the possibility of co-presence in mixed reality. Kim et al. (2018) showed the existence of co-presence between a virtual interlocutor and a human being. Cho et al. (2015) analyzed the influences of co-presence on the performance of virtual staging. They concluded that age and epistemological beliefs significantly influence physical and social presence, as well as the participants' sense of being present together in the CVE. Wienrich et al. (2018) showed that collaborative tasks increase levels of social presence and cooperation in a positive way in large-scale CVE.

In CVE, the base interaction is typically performed by digital avatars, i.e., graphical representations of the users. However, these interactions are limited. Peña Pérez Negrón et al. (2020) concluded that interactions such as interactions verbal/non-verbal, kinesics and proxemics, were used in a positive and discriminating way (Peña Pérez Negrón et al., 2020). Zhang et al. (2018) developed a CVE and a series of collaborative games dedicated to children with autism spectrum disorder, with the aim of establishing the viability and tolerability of the system. They concluded that their system improved collaborative performance among children. Gamelin et al. (2021) confirmed the importance of avatar's fidelity regarding spatial interaction, since it can affect communication between users, especially when performing collaborative spatial tasks.

These studies address the effect that several factors have on collaborative interaction and sense of co-presence; however, most of them explored the digital avatar as the sole cue for inducing co-presence. The videogames industry has devised, implemented, and deployed several other alternatives for entertaining purposes. This paper contributes with a study on the value of these alternatives to amplifying co-presence in virtual house group visits, that is, in a non-entertaining setting.

3. VVAPP DESIGN AND IMPLEMENTATION

3.1 Avatar

A virtual environment is only considered to be a CVE when more than one user is simultaneously present (Casanueva and Blake, 2000) (Buck et al., 2019) (Knapp et al., 2014). Digital avatars are visual demonstrations of the users that contribute for them to feel present in the CVE. The avatar’s visual appearance can be selected to be either stylized or realistic if it clearly represents the user in question. According to Buck et al. (2019), adding just one personalization element to a 3D avatar has more impact of distinguishing than adding several personalization elements. Hence, avatars in VVApp, which have a *stickman*-like look, are distinguished based solely on their colors. Their faces include a nose to cue partners regarding their gazing direction (see Figure 1).

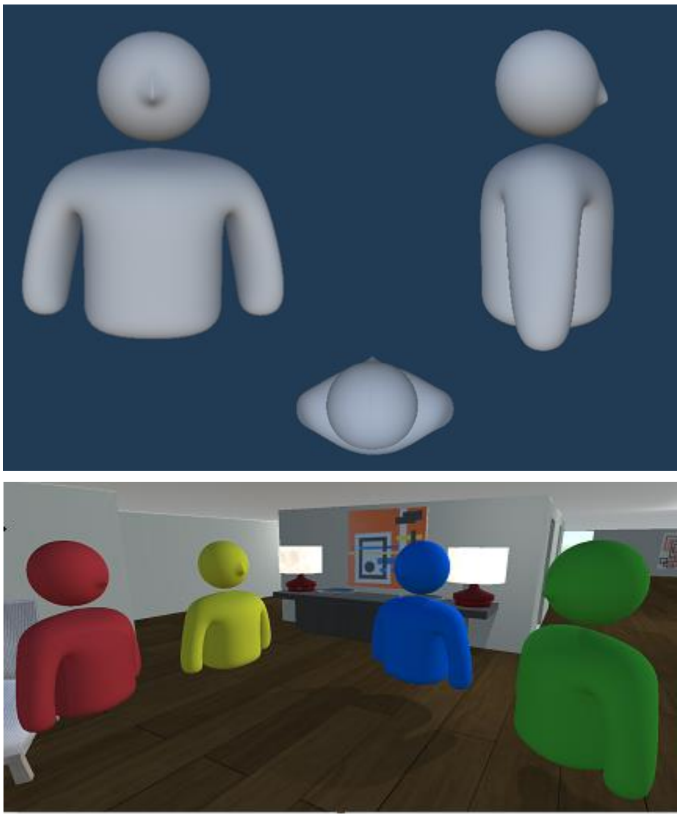


Figure 1. Avatar’s graphical representation (Top). Group of Avatars (Bottom)

To generate interaction events or navigation events within a CVE, a mapping of controls is required. Cursor keys used to be used more in character-based systems, before new alternative control mappings came along outside the keyboard (Dix et al., 2004). In current 3D games, the cursor keys are less used since the inclusion of the mouse. Instead, the WASD mapping is more often used. User controls in VVApp are divided into two parts. The first part follows the pattern used in current video games to control the avatar, of which W moves it forward, S backwards, A to the left, and D to the right. The second part focuses on the mouse position, a variable that modifies the avatar's orientation and the user's perspective, in which the user sees the world in the first person, through the avatar’s eyes.

Hearing reveals significant importance in increasing additional information about a given physical or virtual space. Virtual hearing helps to inform the user the presence and type of sound propagation has occurred. However, it fails to indicate precisely where this sound was propagated in the virtual space. (Török et al., 2015) (Calado et al., 2017). VVApp contains a 3D sound system that fixes this flaw.

3.2 Virtual Houses

The CVE in VVApp presents clues that describe a habitable house so that users feel that they are visiting a real house, such as house rooms, textures to identify the type of flooring or wall texture, and furniture for each room. All elements included in the virtual houses fit together correctly, otherwise the CVE would exhibit unfamiliar visual information, breaking the experience. Currently, VVApp includes three virtual houses of varied sizes, structurally equivalent to real apartments (see Figure 2).

Figure 2. Plan view of the houses CVE_1 (1), CVE_2 (2) and CVE_3 (3). The crosses in the white area indicate the starting position of the four users, i.e., house visitors

House CVE_1 offers a small yet welcoming space, featuring a spacious living room connected to a kitchen, a moderately sized bedroom, and a bathroom. The furnishings are straightforward and functional. In contrast, house CVE_2 presents a medium-sized layout, encompassing a sizable living room, two bedrooms, a bathroom, and a kitchen. This apartment is fully furnished and adorned with decorations. Lastly, house CVE_3 boasts a large floor plan, including a hall path, two bedrooms, a master suite, an office, a spacious living room, an open-space kitchen, a laundry room, a guest bathroom, and a pantry complete with a wine cellar. A broad corridor leads to the undisclosed area, and like in House CVE_2, this apartment is fully furnished and decorated.

3.3 Lobby

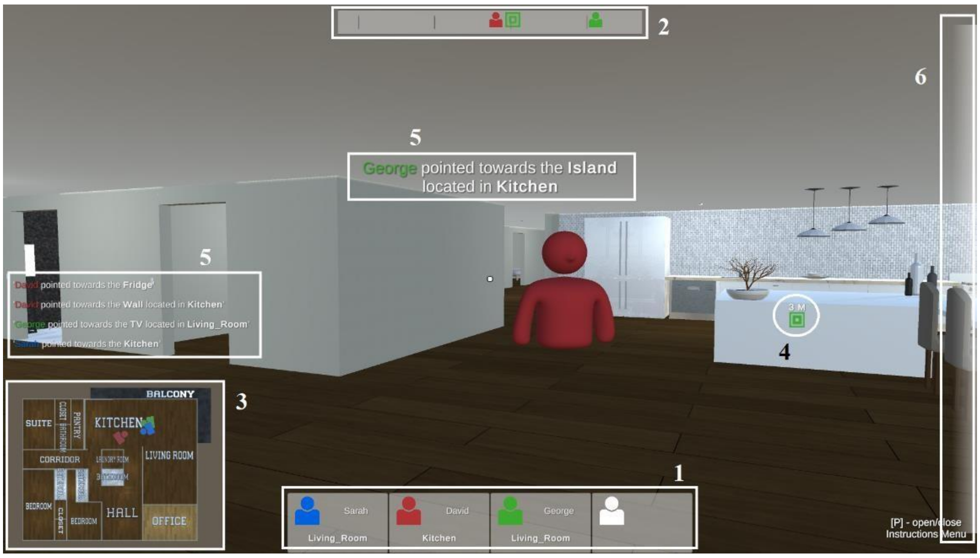
VVApp introduces the user to a functionality called Lobby. This feature allows the user to open a virtual room for others to join, before joining the CVE. The application contains a main menu with options that direct users either to create or join a virtual room. If chosen the first option, the lobby is created, which displays basic information about the CVE (see Figure 3). The user has the option to choose the three virtual houses presented in the Section 3.2. The first iteration of co-presence starts when one or more users joins the lobby, since VVApp notifies their arrival to the first virtual present user, as known as Host. All users have the Ready option, which indicates their availability on exploring the CVE. If all users are ready, VVApp allows the Host to generate the CVE with the virtual house chosen. Many multiplayer video games include this feature as an easy gateway to the sense of co-presence, usually allowing users to interact with each other via verbal interaction functionalities, such as group chat or voice chat.

Figure 3. VVApp’s graphical representation of a Lobby. It is shown the option to change the virtual house (1), the user’s status (2) and the Start Game & Ready buttons (3)

3.4 Collaborative, Non-verbal and Spatial Interaction

Improving the feeling of being together can happen by making sure that essential information is shared among all users in the CVE. Every aspect of the system gives users visual updates about any activity happening in the environment, encouraging a shared understanding and active involvement among everyone involved. (Abrams and Gerber, 2013). Video games, such as *Left For Dead 2* or *Apex Legends*, provide the player with information about their partner’s avatars state. Similarly, VVApp features a rectangular interface *Info-bar* (see Figure 4) segmented into four parts, individually displaying information such as the name and location of each user currently engaged in the CVE. This setup aligns with VVApp's capacity limit, allowing a maximum of four users within the CVE simultaneously.

Video games also often notify players using interfaces about recent events. These events must be clearly notified in advance so that the user experience is not abruptly interrupted. In this line, all users in VVApp are also notified of the others’ information, giving them the feeling that they are not in an empty space (Churchill and Snowdon, 1998). VVApp includes a transparent interface, called *Text box*, which reports the events registered on the CVE (see Figure 4).



George pointed towards the Island located in Kitchen

'David pointed towards the Fridge'

'David pointed towards the Wall located in Kitchen'

'George pointed towards the TV located in Living_Room'

'Sarah pointed towards the Kitchen'

Figure 4. Point of view of Sarah’s avatar. In this screenshot, it is shown all the interfaces added to the application: Info-bar (1), Compass (2), Mini-map (3), Pointing (4), Text boxes (5) and Collision warning (6)

Churchill and Snowdon points out that allowing users to mark three-dimensional positions for others to see (a pointing mechanism) helps collaborative task solving (Churchill and Snowdon, 1998). This reduces communication problems in CVE and potentially amplifies the feeling of co-presence (Gamelin et al., 2021). In this line, the video game *Counter Strike Global Offensive* includes a non-verbal communication tool called *Pointing*, which marks the location defined by a user in three-dimensional space. In VVApp, the *Pointing* mechanism is also featured, and its animated icon includes the relative distance, in meters, between the user and the focus point (Figure 4). Kyriakou et al. (2017) concluded that people feel more present in the virtual environment when there are scenarios with intensive interactions. Supporting this study, videogames such as *Gran Turismo 7*, the player is often notified, via small 2D visual elements, when an opponent approaching from behind is too close. VVApp follows a similar approach by presenting to the user two rectangular graphical elements, by the name of *Collision Warning*, each occupying one of the two sides of the screen (Figure 4). As a partner moves closer to the user within a specific threshold distance, the rectangles gradually become opaquer. However, to ensure clear information transmission, the active graphical element flashes intermittently, disregarding its level of opacity.

Navigation (i.e., planning a path to a desired destination) in large-scale CVE can be challenging for most users if no access to a navigation guide. To address this difficulty, video games often offer players a comprehensive small-scale map of the environment. In CVE’s, maps need to facilitate a user's ability to correlate two points depicted on its graphical representation with their corresponding locations within the actual environment. This correlation is essential for effective navigation within the virtual space. (Darken and Sibert, 1993). VVApp features a rectangular *Mini-map* capturing the entire virtual area (Figure 4). To reduce visual clutter, the decorative objects present in the virtual houses are not rendered on the Mini-map. However, the Mini-map exhibits icons of the avatars present, pinpointing their current position and orientation. In order to enhance the Mini-map, a label representing each house division is superimposed onto the map, displaying the name of that particular division.

Although maps contribute positively for navigation, it is not always the most effective tool (Chen and Stanney, 1999). For instance, videogames such as *The Elder Scrolls V: Skyrim* provides the player with a Compass that, with its compact format, manages to guide equally the player in following the current locations of active objectives or enemies. A similar compass is included in VVApp (Figure 4). It indicates the current orientation of the user's partners, as well as the orientation of the pointings marked with the Pointing tool. The interface of the *Compass* represents points in a 2D format. The *Compass* does not identify the poles that associate with the traditional *Compass*. This feature holds the potential to enhance the sense of co-presence, allowing scenarios such as users guiding their partners by indicating a valid position, and the partner then follows it with the assistance of the Compass, fostering a stronger sense of shared presence within the virtual environment.

3.5 Formative Evaluation

To identify any errors or areas for improvement of VVApp, a formative evaluation session lasting forty minutes was conducted with three remote users. The participants, aged between 22 and 28, consisted of two males and one female. This evaluation involved executing three tasks, these being described in Section 4 within the virtual house CVE_3, during which participants were asked verbal questions about the functionality of the mechanics. Overall, the test yielded positive outcomes concerning the usability of interfaces and non-verbal communication methods. Moreover, all users expressed admiration for VVApp, acknowledging its great functionality concerning co-presence mechanisms and its unique way of originating virtual house tours. All issues with VVApp identified by the participants were rectified. The results presented in the subsequent sections were derived from the updated and improved version of VVApp.

4. EXPERIMENTAL SETUP

To assess the co-presence mechanisms integrated into VVApp, 11 groups, each comprising three participants (without repositioning), were assembled for the experiments. Participants within the same group are henceforth referred to as partners. The participants' ages ranged from 17 to 40 years old, with an average age of 24.03 years and standard deviation $\sigma = 5.02$ (σ and standard deviation will be used interchangeably hereafter). Among the participants, there were 30 males and 3 females. The experiments were designed to evaluate the following working hypotheses:

H1. Participants use at least one of the co-presence inducing interfaces to solve tasks.

H2. The proposed co-presence inducing interfaces give positive collaborative feedback, that is, they are used with some significant frequency by users.

H3. The co-presence inducing interfaces, in general, enrich the visits of virtual houses, that is, they are well received by the users.

The experiments were conducted online with recorded video and sound using NVIDIA GeForce Video Capture software. Participants were informed in advance about the sessions and were required to meet specific technical criteria to access the application remotely, such as using a Laptop or Desktop with at least 8 GB of memory, running Windows or MAC OS, and having downloaded the dedicated version of either Zoom or Discord for verbal communication within their groups. All participants agreed to the session recordings. Each session began with an introductory video tutorial displaying the functionality of VVApp's interfaces from a user's perspective, aiding participants in understanding how to navigate the co-presence inducing features easily. The experiment conductor then explained the tasks and instructed the group to achieve their objectives. After completing a task, participants were asked to leave the server temporarily for data file downloading purposes. The server was brought back online once preparations for the next task were complete, and this process was repeated for each subsequent task. Upon finishing all tasks, groups were provided with a session sheet containing a short questionnaire. The first section covered participants' familiarity with video games, experiences in house-hunting as a customer, and their overall experience using the application. The second section evaluated the perceived usefulness of VVApp's co-presence inducing interfaces through a series of statements, where participants rated their agreement using a 5-point Likert scale ranging from "totally disagree" (1) to "totally agree" (5).

To isolate the analysis of each co-presence inducing interface, a version of VVApp customized for the experiments was developed. In this version, certain interface elements (*interface* hereafter for the sake of compactness) are disabled by default, namely: Interface 1 - *Info-bar*; Interface 2 - *Compass*; Interface 3 - *Mini-map*. To enable them, users must use numerical keys on the keyboard, each corresponding to one of the three interfaces. Enabling one of the three interfaces results in automatic disabling of the other two if they were enabled before. Other interfaces of which were not chosen, remain active regardless of any user’s options. Additionally, VVApp contains a data collection system that saves essential data for the analysis of the experiment, in CSV format. Participants had to execute, in sequence, the following three tasks:

- **Task 1 (Evaluation of virtual houses)** - The group visits the three virtual houses. Participants go through what is necessary to be able to evaluate the house based on their own evaluation criteria (e.g., visual, architectural). The group has a time limit of 4 minutes to navigate inside the house and organize their critique. The evaluation is done individually, that is, each participant must transmit one’s evaluation about the house in question. When the time limit expires, the person in charge of conducting the test asks the participants individually to evaluate the house, even if their assessment is incomplete. During the evaluation, participants can move and use the available co-presence mechanisms. The task ends when the evaluation of the three virtual houses is concluded.
- **Task 2 (Riddle solving)** - The group is asked to solve six riddles in virtual house CVE_3 (see Table 1). The area is large enough to convince participants to explore more cautiously in search for the answers. The answers must be provided collaboratively, that is, all group members must agree with the given decision. The group must answer each riddle within a maximum of 2 minutes and 30 seconds. The group must verbally identify the object they consider to be the solution to the riddle. If they have not reached an answer, the group is given one last opportunity to reach an agreement within a brief period. Participants are free to use all available tools in the application to find the object. The task is completed once the participants answer all the riddles.
- **Task 3 (Pre-defined route)** - The group must follow a predefined route inside virtual house CVE_3. The route contains several intermediate waypoints points that guide participants to their final waypoint. The task starts after the first waypoint is dictated to the participants and they must navigate to that point. This process is repeated from

waypoint to waypoint, until the final waypoint. Participants must stop on the respective waypoint to proceed to the next one. The task has no time limit, nor a mandatory trajectory. Participants have the freedom move as desired between waypoints.

Table 1. Riddles solved by the participants in Task 2

Riddle	Answer
‘You set up, get up and raise your glass daily’	Living room table
‘Well of knowledge, study and also pleasure’	Computer
‘A place we go when we are sad, but also happy’	Pantry wine rack
‘Inform, entertain and decorate’	Television
‘Complicity in a square’	Painting

5. EXPERIMENTAL RESULTS

5.1 Interfaces Utilization Rate

Figure 5 shows the pie charts of the three co-presence inducing interfaces utilization rate. Of the 33 participants, 6 of them did not activate any of three interfaces in Task 1 (18.18%), which contradicts H1. The recordings reveal that these specific participants discussed in more detail the positions of the furniture, the visual appearance of each room, the proportion of the areas, among others, which may justify their lack of need to use to the co-presence interfaces.

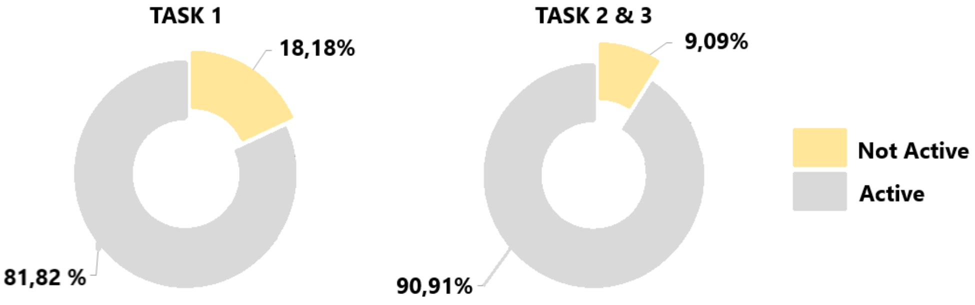
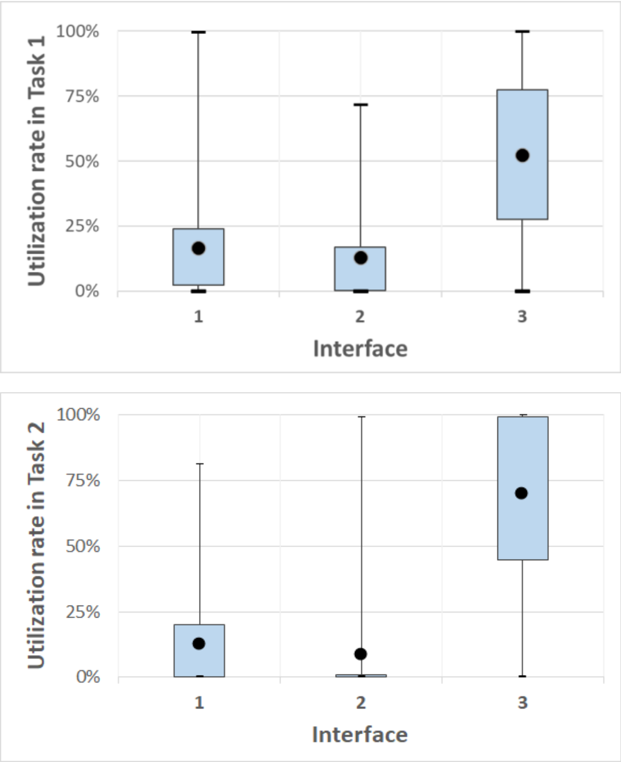


Figure 5. Pie chart of the Interfaces activity

For the remaining participants, a variety of results were obtained. Interface 3 (*Mini-map*) exhibits a utilization rate of 52,22% (σ = 33,15%), which is higher than the 16.56% (σ = 20.91%) obtained for Interface 1 (*Info-bar*) and the 13,04% (σ = 16,03%) for Interface 2 (*Compass*) (Figure 6). It was observed a minimum value of 0 on all interfaces, a maximum value close to 100% for Interfaces 1 and 3, and 72% for Interface 2. Three-quarters of participants utilized less than 23,91% of Interface 1 (σ = 24,00%), less than 16,93% of Interface 2 (σ = 17,00%) and more than 27,58% of Interface 3 (σ = 28,00%). On average, 16.48 transitions between interfaces were recorded (σ = 12.70), which reveals that Interfaces 1 and 2, despite having average utilization rate below 15%, were used occasionally. In Task 2, we can observe differences in relation to the data obtained in Task 1 due to learning effects. Of the 33 participants, 3 of them did not activate any interface, in which 2 of them are the same participants who did not use any of the interfaces in Task 1. These participants were always close to their partners, in the virtual house, while solving riddles. The data files report that participants who did not use any of the interfaces during Task 1, chose Interface 3 (*Mini-map*) in Task 2.

Different results were obtained in Task 2. Interface 3 (*Mini-map*) shows a higher utilization rate compared to Task 1. Interface 1 (*Info-bar*) exhibited a utilization rate of 12,33% (σ = 19,70%), whereas Interface 2 (*Compass*) and Interface 3 (*Mini-map*) exhibited 8,68% (σ = 23,05%) and 69,90% (σ = 36,93%), respectively. It was observed a minimum value of 0 on all interfaces, a maximum value close to 100% for Interfaces 2 and 3, and 81% for Interface 1. Three-quarters of participants utilized less than 20,12% of Interface 1 (σ = 20,00%), less than 00,81% of Interface 2 (σ = 1,00%) and more than 44,66% of Interface 3 (σ = 45,00%). On average, 6.70 transitions between interfaces were recorded (σ = 7.27). Compared with the results of Task 1, there was a decrease of 25.54% in the average utilization rate of Interface 1 (*Info-bar*), a decrease of 33.44% in Interface 2 (*Compass*), and an increase of 25.30% in Interface 3 (*Mini-map*). Although participants resorted more to the *Mini-map*, the number of interface transitions indicate that users used the other two interfaces (Figure 6).



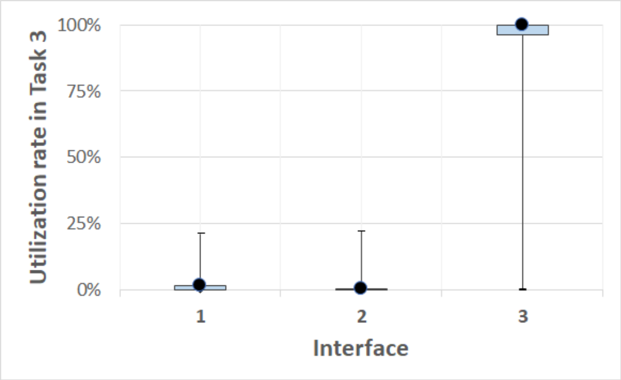


Figure 6. Utilization rates for Interface 1 (*Info-bar*), Interface 2 (*Compass*), and Interface 3 (*Mini-map*) across all tasks. The average value is marked by a circle

In Task 3, we observed differences in relation to the data obtained from Task 2. The same participants reported in Task 2 did not activate any interface. For the remaining, significantly different results were obtained compared to Task 2. Interface 3 (*Mini-map*) once again reveals a higher utilization rate compared to the other two. Interface 1 (*Info-bar*) exhibited an average utilization rate of 2,91% (σ = 6,33\%), Interface 2 (*Compass*) of 1,39% (σ = 4,53%), and Interface 3 (*Mini-map*) of 86,61% (σ = 29,10%). It was observed a minimum value of 0 on all interfaces, a maximum value of 100% for Interface 3, and close to 22% for Interfaces 1 and 2. Three-quarters of participants utilized less than 1,48% of Interface 1 (σ = 2,00%), less than 00,28% of Interface 2 (σ = 0,30%) and more than 96,05% of Interface 3 (σ = 96,00%). On average, 4.52 transitions between interfaces were recorded (σ = 9.28). Compared with the results of Task 2, there was a decrease of 76.40% in the average utilization rate of Interface 1 (info- bar), a decrease of 83.99% in Interface 2 (*Compass*), and an increase of 19.30% in Interface 3 (*Mini-map*). As Task 3 is purely linked to three-dimensional navigation, most participants opted for Interface 3 (*Mini-map*) due to the provided guidance simplicity (Figure 6).

Overall, a varied utilization pattern was noticed among the three interfaces. Interfaces 1 (Info-bar) and 2 (Compass) demonstrated a consistent decrease in usage throughout the entire experiment, while Interface 3 (Mini-map) exhibited an increase in usage. Initially, it may seem like most participants disregarded Interfaces 1 and 2 in the final two tasks. However, the recorded transitions between interfaces tell a different story. Participants utilized Interfaces 1 and 2 sufficiently to gather information about identifying partners within the CVE and their respective positions. The overarching outcomes align with hypothesis H2, supporting the notion that these interfaces offer positive collaborative feedback.

5.2 Collisions and Pointing

Figure 7 presents the number of *collisions* between users and the number of *pointing actions* (to insert a marker in the environment). The main objective of this analysis is to study the participant's behavior from the registered number of collisions and pointing actions. The comparison of values between tasks is not included. In Task 1, a total of 430 collisions and 1343 pointing actions were registered, with an average of 39.09 collisions (σ = 27.79) and of 122.09 pointing actions (σ = 105.46) per group (roughly 1 collision and 4 pointing actions per minute).

Some participants showed care in navigation and curiosity in testing the application in Task 1 (e.g., two of the elements collided on purpose). It was registered 83 collisions and 308 pointing markers as maximum and no collisions or pointing markers as minimum from all participants. Three-quarters of participants collided less than 59 times and pointed more than 36 times. Most participants used pointing actions to identify objects, furniture, and rooms in their speech, which draw the attention of their partners. A total of 295 collisions and 1581 pointing actions were registered in Task 2, with an average of 26.82 collisions (σ = 19.74) and 143.73 pointing actions (σ = 128.98) per group. It was registered 58 collisions and 340 pointing markers as maximum and 3 collisions and a single pointing marker as minimum from all participants. Three-quarters of participants collided less than 43 times and pointed more than 32 times. The participants resorted considerably to pointing actions, as they help users to jointly identify which object may be the correct answer to the riddle. A total of 506 collisions and 311 pointing markers were registered in Task 3, with an average of 46 collisions (σ = 32.27) and 28.73 pointing actions (σ = 27.76) per group. It was registered 108 collisions and 79 pointing markers as maximum and 5 collisions and no pointing markers as minimum from all participants. Three-quarters of participants collided less than 25 times and pointed more than 47 times. Participants collided more frequently in narrower spaces. The pre-defined route included these narrow spaces with the purpose of observing the trajectories chosen by the participants. The *Pointing* mechanism was used to help participants in indicating a clear space when a misunderstanding occurs.

During Task 1, participants extensively utilized pointing markers to provide feedback on house style or furniture placement to their partners during evaluation. They also used these markers to verbally identify objects, rooms, and furniture. Occasionally, collisions between participants occurred. In Task 2, there was evident support indicating that pointing markers significantly aided participants in identifying objects and furniture. The occurrence of collisions notably decreased, suggesting that the collision avoidance mechanism assisted participants in navigating more cautiously. However, Task 3 presented a contrasting scenario with a higher frequency of collisions and a reduced use of pointing markers. This implies that participants' navigation was compromised, particularly in narrow passages. Nevertheless, participants still relied on pointing markers to guide their partners in the correct orientation.

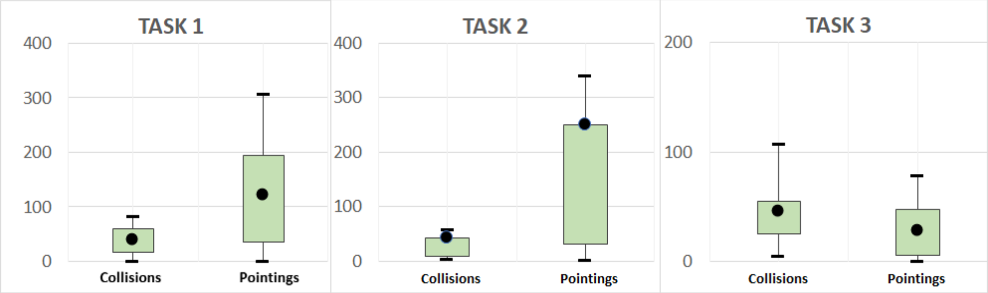


Figure 7. Number of collisions and pointing. The average value is marked by a circle

5.3 Co-presence Mechanisms Subjective Rating

Participants were asked to subjectively rate how often they used each co-presence inducing interface in a 5-point scale (1-not used at all, 5-used very often). Subjective data related to Interfaces 1 (*Info-bar*), 2 (*Compass*), and 3 (*Mini-map*) provided for a given task were discarded for the users that did not actually objectively used any of these interfaces in that task (data in previous sections).

In Task 1, an average rating of 2.19 ($\sigma = 1.04$) was observed for Interface 1 (*Info-bar*), 2.52 ($\sigma = 1.12$) for Interface 2 (*Compass*), 4.74 ($\sigma = 0.59$) for Interface 3 (*Mini-map*), 4.06 ($\sigma = 1.06$) for *Pointing actions*, 3.24 ($\sigma = 1.17$) for *text boxes*, and 2.85 ($\sigma = 1.15$) for the *collision warnings*. Three-quarters of participants rated mechanisms 1 and 2 below 3, mechanism 3 below 4, mechanism 4 at 5 and mechanism 5 above 3 and mechanisms 6 and 7 below 4 (Figure 8). All interfaces, except for the information bar, revealed a good result in terms of usage by the participants. The *Mini-map* and *pointing actions* were consulted with high frequency. Participants reported a neutral position regarding the *text boxes* and the *collision warning*. Some reported that these two mechanisms helped significantly in performing the task, others indicate the opposite effect. Although with lower frequency, participants reported that have consulted the *Compass* occasionally. *Info-bar* was classified as being even less used because participants studied the avatars' appearance right at the task’s onset. In Task 2, an average rating of 2.63 ($\sigma = 1.38$) was observed for Interface 1 (*Info-bar*), 2.53 ($\sigma = 1.17$) for Interface 2 (*Compass*), and 4.53 ($\sigma = 0.97$) for Interface 3 (*Mini-map*), 4.61 ($\sigma = 0.83$) for *Pointing actions*, 2.79 ($\sigma = 1.11$) for the *text boxes*, and 2.24 ($\sigma = 1.09$) for *collision warning*. Three-quarters of participants rated mechanisms 1, 2, 3 and 7 below 3, mechanism 4 above 4, mechanism 5 at 5 and mechanism 6 below 4 (Figure 8). All implemented mechanisms, except the collision warning, revealed a good result in terms of usage by the participants. The *Mini-map* was heavily consulted, and *Pointing* was considered the top choice for riddle solving. Participants reported a neutral position regarding *text boxes*. The *Compass* and *Info-bar* were consulted a little more than in Task 1. The *collision warning* did not help in solving the riddle. In Task 3, an average rating of 2.30 ($\sigma = 1.29$) was observed for Interface 1 (*Info-bar*), 2.17 ($\sigma = 1.12$) for Interface 2 (*Compass*), 4.83 ($\sigma = 0.59$) for Interface 3 (*Mini-map*), 3.03 ($\sigma = 1.51$) for *Pointing*, 2.30 ($\sigma = 1.24$) for the *text boxes*, and 2.76 ($\sigma = 1.32$) for the *collision warning*. Three-quarters of participants rated mechanisms 1, 2 and 6 below 3, mechanism 3 above 4, mechanism 4 at 5 and mechanisms 5 and 7 above 2 (Figure 8). The *Mini-map*, *text boxes* and *collision warning* reveal positive results. The interface of the *Mini map* was considered the top choice for navigation. Participants reported a neutral position regarding *text boxes*. The *Compass* and *Info-bar* were consulted less than in Task 2. The *collision warning*, for few participants, did not help in avoiding collisions, for others, it helped considerably.

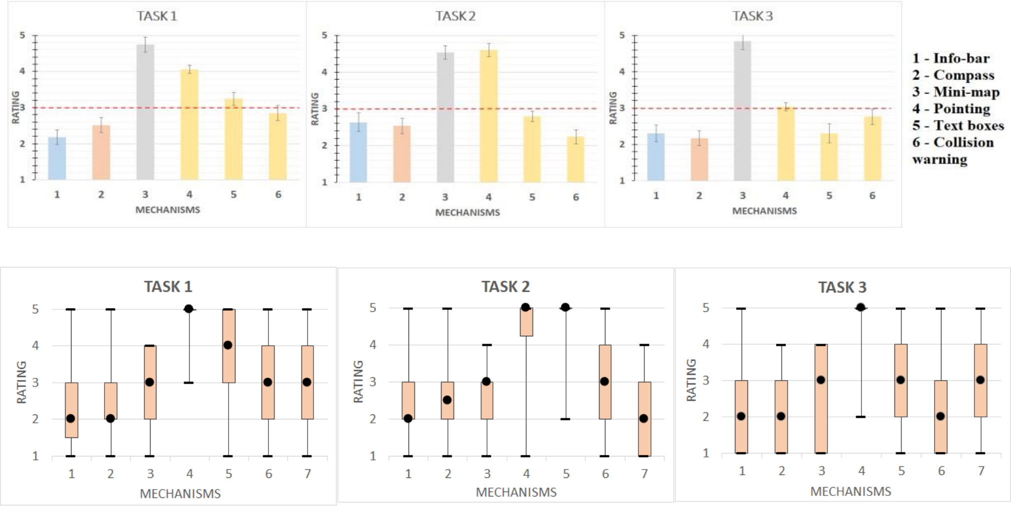


Figure 8. Bar graph of average mechanism rating. It includes a dashed red line that indicates the neutral rating (a value of 3) and the standard error for each mechanism (Top). A box plot of the mechanism rating. The average value is marked by a circle (Bottom)

Participants demonstrated varying preferences for the co-presence inducing interfaces within VVApp across different tasks. In Task 1, the Mini-map emerged as the most frequently utilized tool for navigating inside the virtual houses, while the Info-bar and Compass were sparingly used to locate their partners' avatars and positions. The Pointing feature proved effective as a non-verbal communication tool, participants frequently checked text boxes to identify registered pointing markers, and few paid attention to collision warnings. Moving to Task 2, participants reported using Interfaces 1 (Info-bar) and 2 (Compass) moderately but placed greater emphasis on Interface 3 (Mini-map) for solving riddles. The Pointing mechanism was notably significant in identifying objects and furniture, text boxes were consulted regularly, and collision warnings were not perceived as extremely useful. In Task 3, participants considered the Mini-map as crucial for navigating the virtual house, with the Info-bar and Compass providing intermediate support. The Pointing mechanism was primarily used for guiding participants, while text boxes were consulted regularly. Collision warnings grabbed the attention of few participants, reminding them to navigate cautiously.

This analysis suggests that the implemented co-presence inducing mechanics indeed enhance group visits to virtual houses, as hypothesized in H3. The varied use of interfaces across tasks highlights their utility and contribution to the overall immersive experience during group visits in the virtual environment.

5.4 Times achieved in Tasks 1 and 2

To verify the effectiveness of the co-presence mechanisms implemented on VVApp, Figure 9 shows the execution times and its average for Tasks 1 and 2 in all virtual houses.

Starting with Task 1, in the CVE_1 house, participants took, on average, 3 minutes and 23 seconds (σ = 36.82 s) to evaluate it; in the CVE_2 house, participants took, on average, 3 minutes (σ = 50.83 s); and, in the CVE_3 house, participants took, on average, 3 minutes and 50 seconds (σ = 16.61 s). There is a minimum of 156 seconds for the CVE_1 box, 93 seconds for the CVE_2 box, and 198 seconds for the CVE_3 box; a maximum of 240 seconds for the three spaces. Three-quarters of the participants took more than 166 seconds to evaluate the CVE_1 house, less than 223 seconds to evaluate the CVE_2 house, and more than 225 seconds to evaluate the CVE_3 house.

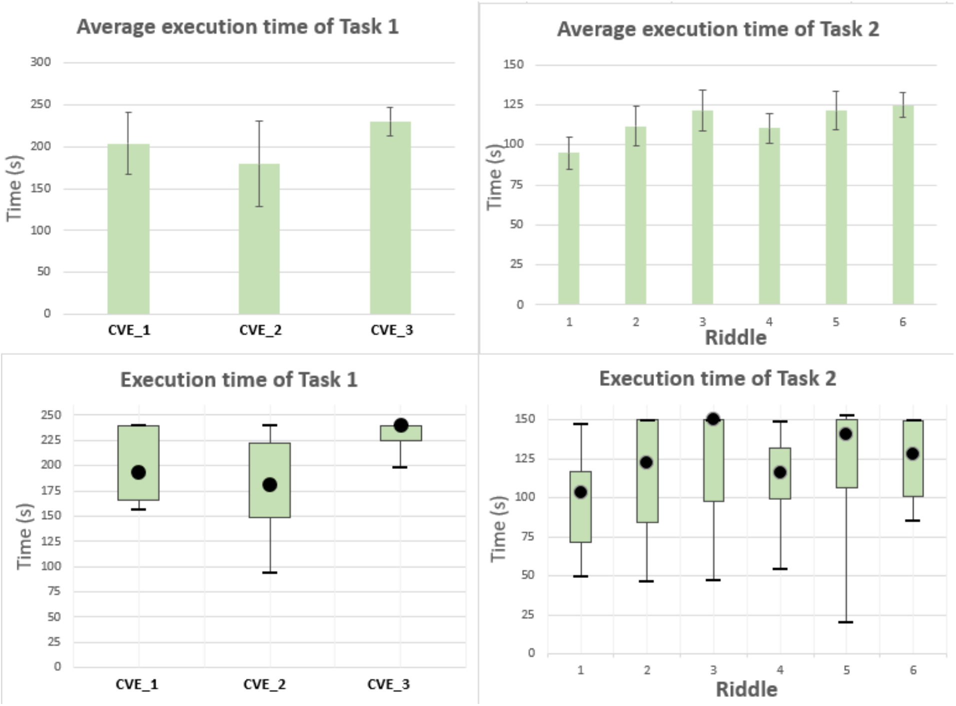


Figure 9. Bar graph of the average execution time in Tasks 1 and 2 (Top). A box plot of the execution time in Tasks 1 and 2 (Bottom)

For Task 2, participants took, on average, 1 minute and 35 seconds (σ = 32.01 s) to solve riddle 1; 1 minute and 52 seconds (σ = 40.26 s) to solve riddle 2; 2 minutes and 1 second (σ = 41.97 s) to solve riddle 3; 1 minute and 50 seconds (σ = 31.08 s) to solve riddle 4; 2 minutes and 1 second (σ = 40.91 s) to solve riddle 5 and 2 minutes and 5 seconds (σ = 25.38 s) to solve riddle 6. It took minimum between 20 and 85 seconds and maximum between 147 and 150 seconds in solving riddles. The recordings and Figure 9 reveal an increase in time throughout the puzzles, since all groups except one got the first puzzle wrong.

Regarding Task 1, the figure indicates a 13.72% decrease in the time taken to complete the task when comparing the first two houses, and a subsequent 28.36% increase when comparing the last two houses. Notably, participants started to increase their collaborative interaction in the application in CVE_1, which is considered the smallest, followed by CVE_2 with a slightly larger area, and finally, CVE_3, the largest. Despite this variation in size, it can be inferred that participants navigated and evaluated houses efficiently across all sizes. Moving to Task 2, a deliberate challenge was introduced in the first riddle to encourage heightened collaborative communication among participants, which effectively occurred. Subsequently, from the second riddle onwards, there was an increased duration in solving the riddles, affording participants ample time to utilize all necessary collaborative interaction mechanisms. Consequently, the groups demonstrated the ability to arrive at the correct answers for certain riddles, highlighting their effective use of collaborative strategies.

5.5 Session Form Questionnaire

Finally, to fill in the session for questionnaire, participants were asked to respond to the following set of four questions:

Q1: “Do you have experience with video games?”
Q2: “Have you ever played co-op video games with partners you know?”
Q3: “Have you been looking for houses to buy/rent before?”
Q4: “Would you consider using a system like the one you've been experiencing to help you choose a house in the future?”

Table 2 summarizes the responses provided by the participants. The responses show that all participants have experience with videogames and recognized the mechanisms implemented in VVApp. All participants, except one, have already played multiplayer videogames with their partners. This reveals that these participants participated in this session with the awareness that they will perform tasks requiring team collaboration. The only participant who had never experienced multiplayer videogames expressed that VVApp’s co-presence mechanisms helped a lot in identifying the group partners. A total of 17 participants have already looked for houses to rent or buy. Of these, 12 are over 21 years old. At the end of the session, all participants praised the collaborative experience.

Table 2. Participants’ responses to the session questionnaire

Questionnaire							
Q1		Q2		Q3		Q4	
Yes	No	Yes	No	Yes	No	Yes	No
33	0	32	1	17	16	33	0

6. CONCLUSION & FUTURE WORK

This article studied the application of video game strategies for user co-presence enhancement in collaborative virtual environments. This study was conducted in the context of virtual group house tours, for which an immersive Unity-based tool was developed, VVApp, targeting the real estate industry and architectural participatory design application scenarios. This tool integrates various videogame-inspired co-presence cues, such as an information bar, a visual compass, a mini-map, pointing actions, text boxes, and collision warnings between users. The results from a user study involving 33 participants, divided in groups of three, yielded positive outcomes. Participants regularly utilized the provided co-presence enhancing features to navigate and communicate more effectively while completing the assigned tasks. Notably, expressed interest in using VVApp in the future.

All participants reported having prior experience with multiplayer video games. Although video games enjoy of high popularity, their use is not universal. Hence, as future work, we plan to assess whether the obtained results hold for people not familiarized with video games, which will surely involve widening the participants' age distribution. This exploration could shed light on the broader applicability and potential enhancements of co-presence features in diverse user demographics. Also, as future work, we also intend to explore and incorporate in VVApp two additional co-presence inducing mechanisms commonly found in video games, these being the text chat and voice chat.

ACKNOWLEDGMENT

This work was partially supported by FCT under ISTAR projects UIDB/04466/2020 and UIDP/04466/2020.

REFERENCES

Abrams, S. and Gerber, H. (2013). Achieving through the Feedback Loop: Videogames, Authentic Assessment, and Meaningful Learning. *The English Journal*, Vol. 103, No. 1, pp. 95-103.

Buck, L., Rieser, J., Narasimham, G. and Bodenheimer, B. (2019). Interpersonal Affordances and Social Dynamics in Collaborative Immersive Virtual Environments: Passing Together Through Apertures. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 25, No. 5, pp. 2123-2133. https://doi.org/10.1109/TVCG.2019.2899232

Calado, J., Santana, P. and Lopes., R. (2017). A unity-based framework for sound transmission and perception in video games. *24º Encontro Português de Computação Gráfica e Interação, Guimarães, Portugal*.

Casanueva, J. and Blake, E. (2000). The Effects of Group Collaboration on Presence in a Collaborative Virtual Environment. In J. Mulder and R. van Liere (eds.) *Virtual Environments 2000. Eurographics*. Springer, Vienna. https://doi.org/10.1007/978-3-7091-6785-4_10

Chen, J. and Stanney, K. (1999). A Theoretical Model of Wayfinding in Virtual Environments: Proposed Strategies for Navigational Aiding. *Presence*, Vol. 8, No. 6, pp. 671-685. https://doi.org/10.1162/105474699566558

Cho, Y. H., Yim, S. Y. and Paik, S. (2015). Physical and social presence in 3d virtual role-play for pre-service teachers. *Internet and Higher Education*, Vol. 25, pp. 70-77. https://doi.org/10.1016/j.iheduc.2015.01.002

Churchill, E. and Snowdon, D. (1998). Collaborative virtual environments: an introductory review of issues and systems. *Virtual Reality*, Vol. 3, *No.* 1, pp. 3-15. https://doi.org/10.1007/BF01409793

Darken, R. and Sibert, J. (1993). A toolset for navigation in virtual environments. *Proceedings of the 6th annual ACM symposium on User interface software and technology*, pp. 157-165.

Dix, A., Finlay, J., Abowd, G. and Beale, R. (2004). *Human-computer interaction*. Pearson Education.

Fernandes, J., Brandão, T., Almeida, S. M. and Santana, P. (2023). An educational game to teach children about air quality using augmented reality and tangible interaction with sensors. *International Journal of Environmental Research and Public Health*, Vol. 20, No. 5, 3814.

Gamelin, G. et al. (2021). Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments. *Pers Ubiquit Comput*, Vol. 25, pp. 467-484. https://doi.org/10.1007/
s00779-020-01431-1

Kim, K., Bruder, G. and Welch, G. (2018). Blowing in the Wind: Increasing Copresence with a Virtual Human via Airflow Influence in Augmented Reality. *International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*. Retrieved from https://par.nsf.gov/biblio/10105838. https://doi.org/10.2312/egve.20181332

Knapp, M., Hall, J. and Horgan, T. (2014). *Nonverbal Communication in Human Interaction*. Wadsworth, Cencage Learning.

Kyriakou, M., Pan, X. and Chrysanthou, Y. (2017). Interaction with Virtual Crowd in Immersive and semi- Immersive Virtual Reality systems. *Computer animations virtual worlds*, ISSN 1546-4261, Vol. 28, Issue 5. https://doi.org/10.1002/cav.1729

Peña Pérez Negrón, A., Muñoz, E. and Lara López, G. (2020). A model for nonverbal interaction cues in collaborative virtual environments. *Virtual Reality*, Vol. 24, pp. 605-618. https://doi.org/10.1007/s10055-019-00421-w

Pimentel, D. and Vinkers, C. (2021). Copresence with virtual humans in mixed reality: The impact of contextual responsiveness on social perceptions. *Frontiers in Robotics and AI*, Vol. 8, 634520. https://doi.org/10.3389/frobt.2021.634520

Podkosova, I. and Kaufmann, H. (2018). Mutual collision avoidance during walking in real and collaborative virtual environments. *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D '18)*. Association for Computing Machinery, Article 9, pp. 1-9. https://doi.org/10.1145/3190834.3190845

Stanney, K. and Cohn, J. (2009). Virtual environments. *Human-Computer Interaction*, pp. 1031-1048.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, Vol. 42, Issue 4, pp. 73-93.

Török, Á. et al. (2015). It sounds real when you see it. Realistic sound source simulation in multimodal virtual environments. *Journal on Multimodal User Interfaces*, Vol. 9, pp. 323-331. https://doi.org/10.1007/s12193-015-0185-4

Vinson, N. G. (1999). Design guidelines for landmarks to support navigation in virtual environments. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '99). Association for Computing Machinery, pp. 278–285. https://doi.org/10.1145/302979.303062

Warpefelt, H. (2016). *The Non-Player Character: Exploring the believability of NPC presentation and behavior* (PhD dissertation). Department of Computer and Systems Sciences, Stockholm University.

Wienrich, C., Schindler, K., Döllinqer, N., Kock, S. and Traupe, O. (2018). Social Presence and Cooperation in Large-Scale Multi-User Virtual Reality - The Relevance of Social Interdependence for Location-Based Environments. *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 207-214. https://doi.org/10.1109/VR.2018.8446575

Zhang, L. et al. (2018). Understanding Performance and Verbal-Communication of Children with ASD in a Collaborative Virtual Environment. *Journal of Autism and Developmental Disorders*, Vol. 48, pp. 2779-2789. https://doi.org/10.1007/s10803-018-3544-7