

Interaction with Virtual Crowd in Immersive and semi-Immersive Virtual Reality systems

Marios Kyriakou, Xueni Pan, Yiorgos Chrysanthou

This study examines attributes of virtual human behavior that may increase the plausibility of a simulated crowd and affect the user's experience in Virtual Reality. Purpose-developed experiments in both Immersive and semi-Immersive Virtual Reality systems queried the impact of collision and basic interaction between real-users and the virtual crowd and their effect on the apparent realism and ease of navigation within VR. Participants' behavior and subjective measurements indicated that facilitating collision avoidance between the user and the virtual crowd makes the virtual characters, the environment and the whole Virtual Reality system appear more realistic and lifelike. Adding basic social interaction, such as verbal salutations, gaze, and other gestures by the virtual characters towards the user, further contributes to this effect, with the participants reporting a stronger sense of presence. On the other hand, enabling collision avoidance on its own produces a reduced feeling of comfort and ease of navigation in VR. Objective measurements showed another interesting finding that collision avoidance may reduce the user's performance regarding their primary goal (navigating in VR following someone) and that this performance is further reduced when both collision avoidance and social interaction are facilitated.

1. INTRODUCTION

Virtual humans are commonly used in Virtual Reality (VR) applications either as part of the environment or as the main interface for the user. Research studies have demonstrated that participants would react and behave towards virtual humans in a highly realistic manner in terms of their subjective, objective, and physiological measurements [1], [2]. Despite the fact that consciously participants are aware that virtual humans are computer generated agents, they would go as far as to automatically attribute mental states to them [3]. Understanding how the participant is being influenced and how they react towards virtual humans could help us develop VR applications that are more convincing and more immersive.

Many experimental studies have been conducted in the area of understanding the interaction between virtual humans and human participants, with the intention of pinning down specific characteristics of the verbal or non-verbal behavior of the virtual humans that would influence directly the participants' behavior. When participants are placed in VR with a virtual crowd, where the environment is populated with a relatively big groups of virtual humans spatially distributed with different behaviors (walking alone or in groups, staring at shop windows, talking, waiting etc.), there are more things to consider regarding how participants' experience and behavior are being influenced by their interaction with the virtual crowd. Researchers using experiments explored the impact of characteristics of groups in the received realism of virtual crowds in VR. Still, for Immersive Virtual Reality (IVR) systems, there is no sufficient research on how different levels of interaction between the participant and the virtual crowd affects the participant's level of presence, or in other words the realism of the virtual crowd's action and the participants' subjective perception of realism [4]. In the following we review the existing research on participants' behavioral reaction towards virtual humans and how virtual crowd is perceived in VR.

2. RELATED WORK

Among the research studies with individuals or small groups of virtual humans, many have focused on the participants' behavior in maintaining the interpersonal distance with virtual humans (proxemics). Participants tend to maintain a greater distance with more realistic agents [5] and show negative reactions to violations of interpersonal space [6]. In other studies in IVR, with up to 4 virtual humans [7], [8], a few interesting outcomes were reported regarding the distances that participants maintain with virtual humans, how they were defined and governed: (1) participants showed increased physiological reaction the closer they are approached by virtual characters; (2) participants maintained greater distance from virtual humans when approaching their fronts compared to their backs; (3) participants gave more personal space to virtual characters who engaged them in mutual gaze; and (4) participants moved farthest from virtual humans who entered their personal space. Also, experiments in a recent study [9] have shown that participants tend to go around dense groups of virtual humans and through sparse ones. These motivated us to study further how collisions between virtual humans and the participants affect the participants' behavior.

A number of studies have been carried out investigating how we perceive virtual crowds in VR. In experimental studies [10]–[12], researchers examined the impact of characteristics of groups in the perceived realism. They found that the addition of groups of virtual humans improved the realism of crowd scenes if the group sizes and numbers were plausible. Researchers have studied the effects of the positions and the orientations of the virtual humans on the plausibility of the crowd, concluding that rule based crowd formations are more realistic than random formations [13]. Studying the participants' perception of virtual crowd's realism, Ennis et al. [14] found that one of the most important factors is the context of the scene and whether virtual humans' behavior was adherent to the context.

When it comes to the measurement and evaluation of VR, the interaction with virtual humans has always been recognized as an important part in the level of presence reported by the participants. Schubert et al. [15] defined interaction as “manipulation of objects and the influence on agents”. Garau et al. [16] found that participants' level of presence diminished if the virtual humans in the scene were not interactive [17]. There is also compelling evidence that the interaction between human participants and individual virtual humans can be highly realistic [3]. In particular, Slater et al. [18] found that the users' heart rate increased when a virtual human talked to them. Moreover, Pelechano et al. [19] explored the egocentric features that a virtual crowd should have in order to achieve high levels of presence. They discovered that breaks in presence or low comfort appeared when there was no collision avoidance and when there were shaking artifacts in the crowd simulation method. Similar conclusions were also extracted in the study of Ahn et al. [20]. Thus, a significant factor that has an effect on users is their interaction with virtual humans.

Many aspects concerning the relationship between the user and the virtual crowd in a VR system remains to be studied. The objective of our study is to investigate the user's behavior and their subjective perception of realism, concentrating on the degree of interaction between the user and the virtual crowd, focusing on two main crowd behavior characteristics: collision avoidance, and basic social interaction (such as salutation and gaze) between the virtual crowd and the participant. Here we include collision avoidance because it has been the main concerns of the crowd navigation methods that generate the low-level crowd behavior [21], [22]. Also, in a recent study [23], it was found that users acted differently (decrease of walking speed, increase of the clearance distance) when they had to avoid collisions with virtual humans than with other inanimate objects. Our hypothesis is that the level of interactivity of a virtual crowd towards the immersed participants would have an impact on participants' felt level of presence and their performance in VR.

In particular, we ran experimental studies in two systems: a CAVE-like system (IVR) where participants navigate via a wand, and a single wall system (semi-IVR) where participants navigate using the walk-in-place paradigm (via a Kinect). The main purpose of this study, however, is not to compare the two systems, but rather testing our hypothesis on virtual crowd interaction in our two existing systems. This work is the extension of [24] where preliminary results showed that, with our setup, it is more important for our walk-in-place semi-IVR system to facilitate collision avoidance between the user and the virtual agents accompanied with basic interaction between them. Findings from this paper could help future development in generating realistic looking virtual crowds by avoid demanding work on characteristics that contribute less on the virtual crowd and concentrate on more important ones that enhance the user's sense of presence.

3. METHODS

The approval for the experiments was granted by the National Bioethics Committee. Two experimental studies were conducted: one with twenty volunteers in IVR system experiments and another one with thirty volunteers in semi-IVR system experiments. Each volunteer participated in three different scenarios, resulting in 60 single-user different sessions in IVR (20 participants) and 90 in semi-IVR (30 participants). The design of the experiment was repeated-measures (within-subjects), testing all participants under different levels of interaction. An important reason for using a repeated-measures design is that it introduces less variance due to participant disposition [25]. Participants who are prone to being scrupulous will likely exhibit the same behavior in all the experiments they will participate in. Thus, the variability of the experimental results is more dependent on the different levels of interaction, rather than on behavioral differences between participants. In the following section, we describe the technical setup, the scenario, and the procedure and measurements of the experiment.

3.1. The system

The 3D interactive virtual environment was developed using the Unity3D¹ game engine. Several virtual character models were used in the scenarios, featuring different faces and somatotypes. The animations of the virtual agents were motion-captured offline. A volunteer was asked to perform several different motions, which were recorded using a Phasespace Impulse X2 system and manipulated in Autodesk's MotionBuilder² prior to importing them into the Unity3D game engine. Motions were semantically segmented (i.e. walk, turn, stand, talk, wave, etc.) and were programmatically used in the scenarios. This allowed us to synthesize complex and dynamic behaviors for virtual characters in real-time.

The experiments took place in two different VR systems: an immersive and a semi-immersive one. The study does not try to evaluate the effect of the system, but important factors in terms of crowd behavior towards the user's experience.

3.1.1. Immersive VR System

The first set of experiments was conducted in a Cave-like projection based system [26]. This has three back-projected vertical screens (front, left and right) (3m×2.2m) and a floor screen (from a ceiling mounted projector) (3m×3m). Participants' heads were tracked with an

¹<http://www.unity3d.com>

²<http://www.autodesk.com/products/motionbuilder/overview>

InterSense IS 900 tracker, and the navigation through the environment was facilitated through a wand (Figure 1).



Figure 1: A participant using the wand to navigate in the CAVE.

3.1.2. Semi-Immersive VR System

A custom-built semi-Immersive VR system was used for the second set of experiments, using a large screen front-projected wall, driven by a workstation computer with an Intel Pentium i5 3.2Ghz CPU, 8GB of RAM and an NVidia GeForce 525M graphic card.

Using a Kinect [27] for motion sensing and human body tracking, the participants were able to navigate into the virtual world. In order to move forward in the virtual world, the participants walked in place (Figure 2). To rotate their view, they raised their arm in the height of their shoulder (Figure 3) (left hand for rotating to the left; right hand to rotate to the right). The participants could walk and rotate at the same time (Figure 4).



Figure 2: A participant walks in place to move forward in the virtual world.

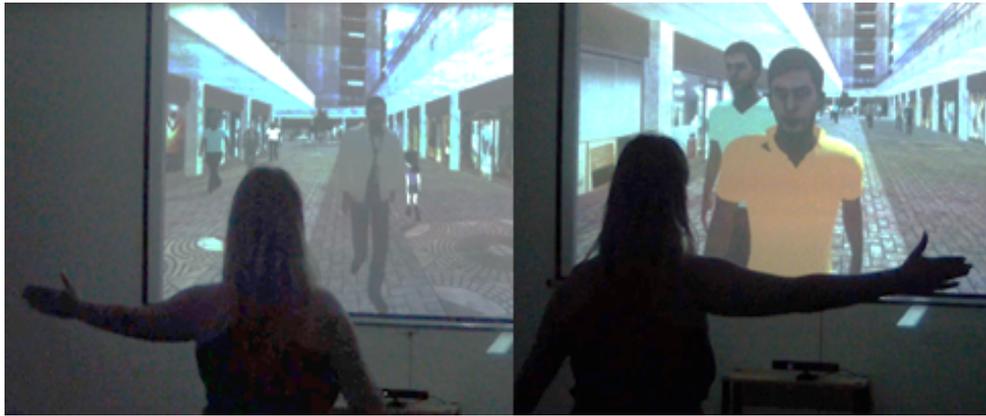


Figure 3: A participant raises her left arm to rotate to the left and her right arm to rotate to the right.



Figure 4: A participant walks and rotates at the same time.

3.2. Scenarios

We designed a 3D virtual environment representing an open-space mall with a significant number (33) of animated virtual characters, spatially distributed with different appearances and behaviors forming a crowd of virtual characters. All virtual characters were programmed with collision avoidance behavior and some basic interact-with-user predefined behaviors, which required no intervention by an operator but were enabled only in the third scenario (as described below - S3).

The participants were given the task to locate a child (a little girl) and follow her wherever she went. This was their primary goal and was clearly stated to them. In particular, the participants were told to try to be at a close distance to the child at all times, navigating into the virtual world. The virtual child was programmed to be singing loudly while following a trajectory, where she came across other virtual characters, mostly coming from the opposite direction, and avoiding collisions with them (Figure 5).



Figure 5: Following a virtual child (little girl) going in the opposite direction of a group of other virtual characters.

The trajectories of the virtual characters were preprogrammed in a way that brought many of them face-to-face with the participant. The purpose of this was to have several possible interaction points between the participant and the virtual characters.

We distinguished three levels of interaction between the virtual crowd and the user. Based on this, we designed three different scenarios, introducing in each one a different level of interaction:

Scenario 1 (S1): the virtual crowd ignores the participant (the virtual characters do not avoid any collision with the participant, and have no other interaction with him/her) (Figure 6).

Scenario 2 (S2): the crowd avoids collisions with the participant but has no other interaction (Figure 7).

Scenario 3 (S3): the crowd interacts with the participant using some basic social interaction (talking to him, looking at him, waving etc.) as well as applying collision avoidance with the participant (Figure 8). All virtual characters avoid collisions with the participant, but only a small number of them actually interact with him/her in more ways.

The order that the three scenarios were presented to the participant was random, so as to get a more objective feedback.



Figure 6: Scenario 1 - virtual crowd ignores the participant without even trying to avoid collisions.



Figure 7: Scenario 2 - virtual crowd avoids any collisions with the participant.



Figure 8: Scenario 3 - virtual crowd interacts and avoids collisions with the participant.

3.3 Procedure and measurements

All participants were informed regarding the procedures of the experiment, the equipment they would use and were informed that they could withdraw from the experiments at any time. They also gave their permission to be filmed. Finally, they completed a three-minute training session using the VR system (the IVR or the semi-IVR) and navigation system prior to the actual experiment, in order to familiarize with the system. During the training session, no virtual human was present.

After each scenario, the participants were asked to fill in a web-based questionnaire. The first questions concerned their gender and their prior experience with video games, while the rest of the questions addressed their experience for the scenario they had just completed. The post-experiment questionnaire were based on the presence questionnaire of Witmer and Singer [28]. Some questions were about the virtual human's awareness of each other and of the participant's presence, while others asked about the realism of the virtual characters and the environment. There were also questions about the participant's comfort, sense of presence and the ease of completing the task.

Alongside questionnaires, we also collected participants' behavioral responses. As indicated in the literature, when studying presence, questionnaires are not viable as the only means for receiving participants' feedback. Experts suggest using both subjective and objective methods [29], [30]. More specific, researchers suggest that *task performance* can be used as an objective measure of presence [31], [32]. Thus, as a more objective method, we used the trajectory analysis which is the result of the task performance of the participants. The task for the participants was to follow a child, since this gave them the ability to move in the virtual environment, come across with virtual characters and interact in various ways with them. Moreover, we had the means to study how well the participants performed, i.e. follow closely the child, since this was their task. During each experiment, the trajectories in the virtual world

of each participant and the virtual characters were recorded. Our main analysis interest about the task performance was in the distance between the participant and the child in the virtual world during the experiment, calculating how close and for how long the participant remained with the child. This was used as a goal achieving evaluation.

4. RESULTS

As mentioned in 3.3, participants answered a questionnaire (see supplemental material) with eight closed-ended questions on a Likert scale ranging from 1 to 5 (1=Minimum, 5=Maximum). In order to understand the participants' responses, we categorize our questionnaire results into 3 different categories: Validation (*Aware_self*) as a check for the validity of the participants' answers, Presence (*Aware_others*, *Presence*, *Realism_Child*, *Realism_Crowd*, and *Realism_Env*) as questions concerning the user's sense of presence, and Performance (*Easiness* and *Comfort*) as questions which reveals their subjective level of ability to complete their task. Finally, we present the results from our behavior measurements (distance analysis).

Question Coding	Question Description	Category
<i>Aware_self</i>	Virtual characters aware of myself	Validation
<i>Aware_others</i>	Virtual characters aware of each other	Presence
<i>Presence</i>	Feeling of presence	Presence
<i>Realism_Child</i>	Realism of child	Presence
<i>Realism_Crowd</i>	Realism of the virtual crowd (except for child)	Presence
<i>Realism_Env</i>	Realism of environment	Presence
<i>Easiness</i>	Easiness of following the child	Performance
<i>Comfort</i>	Feeling comfortable	Performance

Table 1: Questions descriptions and categories.

4.1. Reliability and Validity

To ensure that the internal reliability and validity of the categories met the minimum requirements, we assessed the reliability of the categories with Cronbach's alpha. The results showed that Cronbach's alpha was above the recommended one (Cronbach's alphas for the 5 Presence and 2 Performance questions were 0.84 and 0.86 respectively - the category Validation has only one item). Thus, it can be concluded that the categories have internal consistency and reliability.

4.2. Validation question

This section concerns the validation of our method and manipulation. We used the first question (*Aware_self*), which concerned the crowd's awareness of the participant, as a validation check for the overall participants' responses. Our hypothesis is that in S3 the crowd's awareness of the participant would be stated as the highest, while the lowest one would be stated in S1. As expected, the perceived virtual crowd's awareness of themselves was significantly different between the three scenarios in both IVR and semi-IVR. The Wilcoxon signed-rank test further suggested that, for both IVR and semi-IVR, the level of the crowd's awareness was higher in S2 compared with the level in S1, higher in S3 than in S1, and higher in S3 than in S2 (all test show IVR: $p < 0.001$; semi-IVR: $p < 0.001$) (see supplemental material Tables 7 - 13).

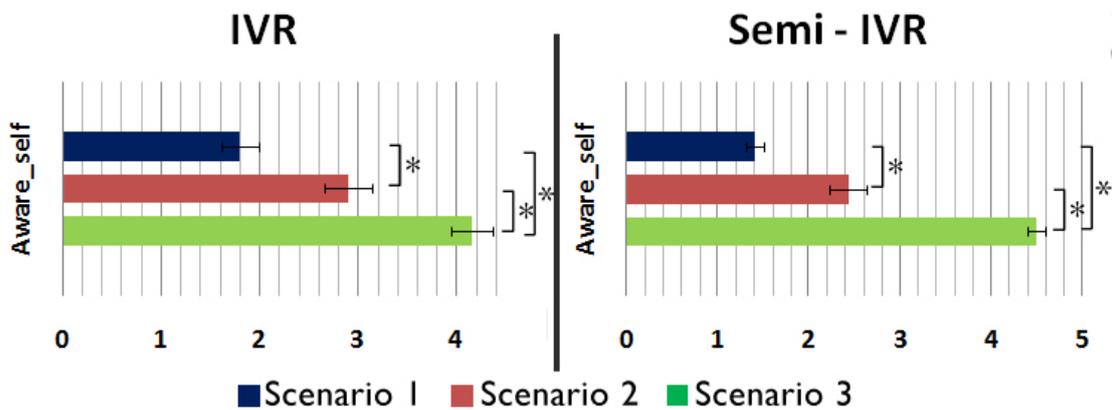


Figure 9: Evaluation of awareness of myself (*Aware_self*). Means of participants' answers of both systems. Error bars present standard error of means. * = $p < 0.001$.

4.3. Presence

We present the results in Figure 10. For more details please see supplemental material table 7-13. Here we present a summary of our findings for each variable in terms of participants' reported level of *Presence* and their perceived level of realism of our environment for each scenario. Please note that the level of realism for all aspects presented here in fact remains consistent in all three scenarios. We predict an increased level of *Presence* from S3 to S1 ($S3_Presence > S2_Presence > S1_Presence$), and the same pattern due to their perception of realism caused by increased level of presence.

Examining the answers of *Aware_others* gave us some interesting findings. The awareness among the virtual characters was programmed to be at the same level across the three scenarios. Still, the participants falsely believed that it had been raised from S1 (no collision avoidance and no interaction between the participant and the virtual crowd) to S2 (with collision avoidance enabled but no other interaction). This belief was even stronger in S3 (with both collision avoidance and interaction enabled) in the semi-IVR system. The difference between the 3 scenarios of was significant for semi-IVR but not for IVR. The Wilcoxon signed-rank test revealed that the level of *Aware_others* was significantly higher in S2 than in S1 for both systems. Also, in S3 it was significant higher than in S1 for the semi-IVR system but not for the IVR, and similar result was obtained between S3 to S2.

The stated level of the *Presence* feeling was significantly different between the 3 scenarios in both IVR and semi-IVR systems. As expected, the Wilcoxon signed-rank test suggested that, for both systems, this level was significantly higher in S2 than in S1, higher in S3 than in S1, and higher in S3 than in S2.

The question *Realism_Child* addressed the perceived realism of the child. Note that the participant was almost always behind the child, trying to catch up with it and there were almost no collisions or interactions between the participant and the child. The differences between the 3 scenarios were statistically significant in both IVR and semi-IVR systems. The Wilcoxon signed-rank test suggested that, in IVR, the level of the perceived realism of the child was significantly higher in S2 than in S1 and higher in S3 than in S1. The perceived realism of child was also higher in semi-IVR but not significant. There was no statistically significant difference between S3 and S2 in either system.

The realism of the crowd (*Realism_Crowd*) was stated as significantly different in both systems. As expected, the Wilcoxon signed-rank test further suggested that, for both IVR and semi-IVR, the level of the crowd realism was significantly higher in S2 than in S1, higher in S3 than in S1, and higher in S3 than in S2.

The virtual environment was exactly the same in all three scenarios. Nevertheless, answers to question *Realism_Env* exhibited a slightly higher score for S2 than in S1 for both IVR and semi-IVR. The difference between the 3 scenarios was statistically significant only in the semi-IVR. The Wilcoxon signed-rank test suggested that, for semi-IVR, *Realism_Env* was significantly higher in S2 than in S1 and in S3 compared to S1, but not between S3 and S2.

Overall, it is clear that in both IVR and semi-IVR systems that S3 triggered the highest level of presence across most measurements and S1 the lowest. This is consistent with our hypothesis. The effect is in particular strong in semi-IVR where all 5 measurements revealed the same pattern. In IVR, 3 out of 5 measurements pointed towards the same direction. The other two measurements in IVR (*Aware_others* and *Realism_environment*) suggested that S1 still was the lowest level, but S3 not the highest.

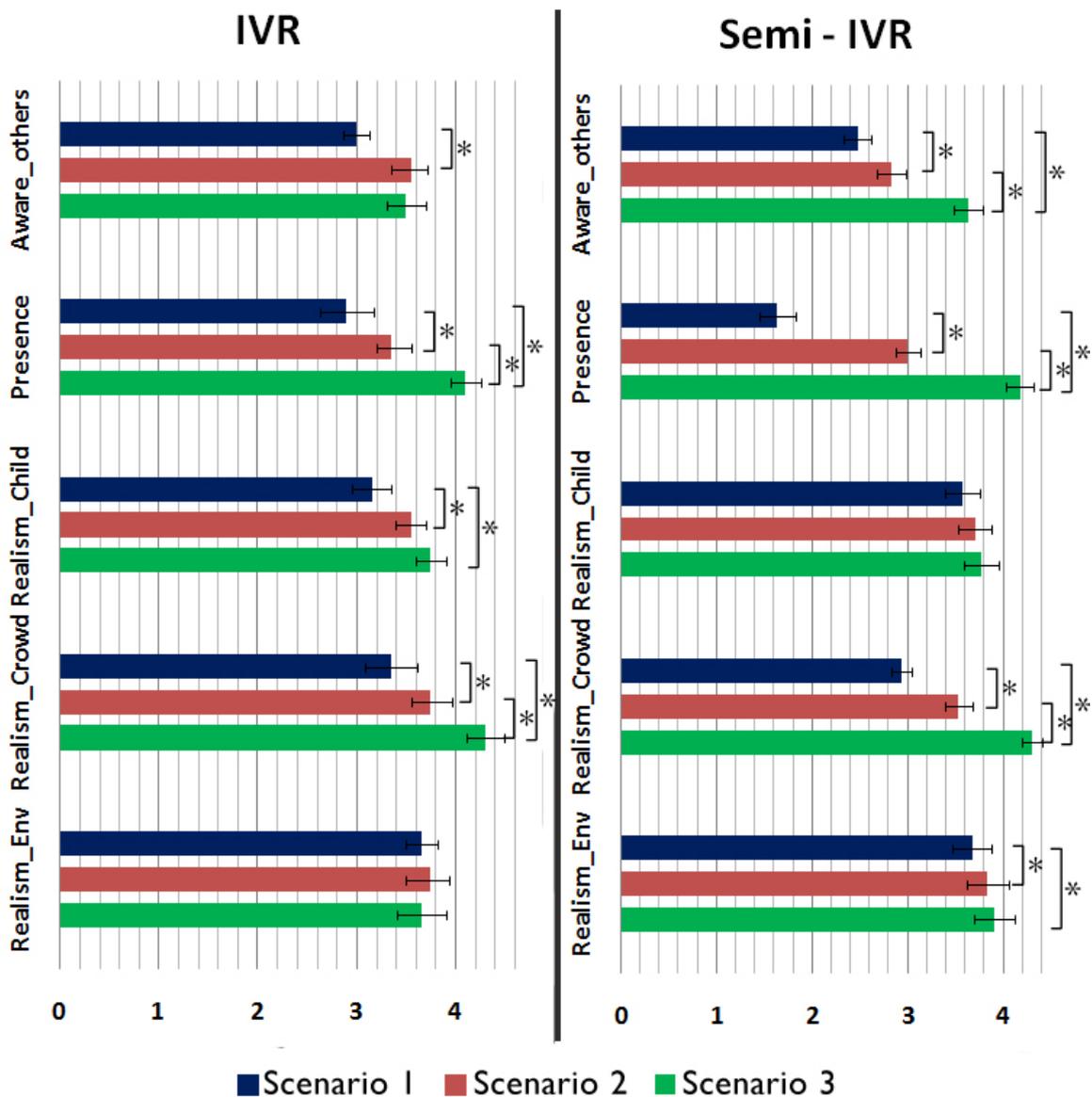


Figure 10: Evaluation of feeling presence questions of both systems. Error bars present standard error of means. *: $p < 0.05$ (see supplemental material Tables 7 - 13).

4.4. Subjective performance- Goal Achievement

Two questions were included regarding participants' *Performance*: (1) *Easiness*: how easy it was for them to complete their target (follow the child) and (2) *comfort*: how comfortable was the use of the system. Our results here suggested an interesting pattern.

Easiness was significantly different in all scenarios in both systems. The Wilcoxon signed-rank test further suggested that, for both IVR and semi-IVR, the level of *Easiness* was the lowest in S2 (significantly lower than both S1 and S3), and the highest in S3 (significantly higher than both S1 and S2).

Comfort, had an identical pattern as *Easiness*: it was significantly different in all scenarios in both IVR and semi-IVR system, and the Wilcoxon signed-rank test revealed S2 has the lowest and S3 the highest.

To summarize, participants found the S3 the easiest and most comfortable scenario compared with the other two in both IVR and semi-IVR, and S2 the hardest and least comfortable one. Moreover, there was a significant positive correlation between the ease of following the child (*Easiness*) and the feeling comfort (*Comfort*) for both systems (Pearson's: IVR $r = 0.47$, $p < 0.05$ and semi-IVR $r = 0.48$, $p < 0.05$). Different from the pattern of participants' perceived level of realism and presence ($S3_Realism > S2_Realism > S1_Realism$), here in terms of their perceived level of performance we have a very different but interesting pattern ($S3_Performance > S1_Performance > S2_Performance$).

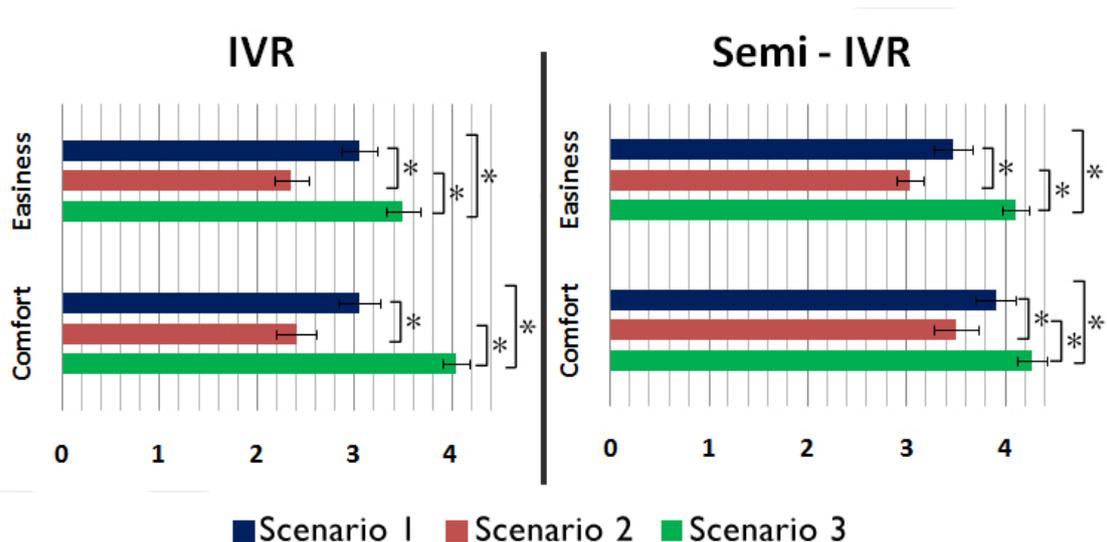


Figure 11: Evaluation of ease of following the child (*Easiness*), and feeling comfort in the VR system (*Comfort*) of both systems. Error bars present standard error of means. * = $p < 0.05$ (see supplemental material Tables 7 - 13).

4.5. Behavioral Analysis

Other than perceived performance, we also look at participants' actual performance as these two often reveals very different patterns.

During the experiment, the trajectories in the virtual world of the participant and the virtual characters were recorded and analyzed. From the trajectories over time we extracted objective measurements for the participants' performance. In particular, participants were told that their

goal was only to follow the child that was in front of them and remain close to it wherever it went. We concentrated our analysis on the distance between the participant and the child during the experiment, measuring how close and for how long the participant was to the child. More specifically, we took three measurements (in meters) and calculated their averages for each scenario: the minimum, maximum and average distance. The distances were measured from the center of the child to the center of the participant.

In addition, we calculated the time (in seconds) that the participant remained more than a certain distance away from the child (Table 2). We set up an arbitrary cut-off distance of five meters ($T_{D>5}$) for all experiments. The duration of the experiments was always the same. The child was following the same trajectory with the same speed, and each experiment was terminated when the child reached the end of her trajectory.

Variable	Description
D_{min}	The minimum distance (in meters) between the participant and the child during the experiment.
D_{max}	The maximum distance (in meters) between the participant and the child during the experiment.
D_{avg}	The average distance (in meters) between the participant and the child during the experiment.
$T_{D>5}$	The time (in seconds) that the participant remained more than five meters away from the child.

Table 2: Variables - Objective Analysis.

Inspecting the following figures (Figure 12 and Figure 13); we can infer that the participant managed to be closer to the child in S1 when there was no collision avoidance and no interaction between the participant and the virtual crowd. The worst scores were recorded in S3 with collision avoidance and basic interaction enabled. The second scenario, with collision avoidance enabled but no other interaction, was somewhere in the middle.

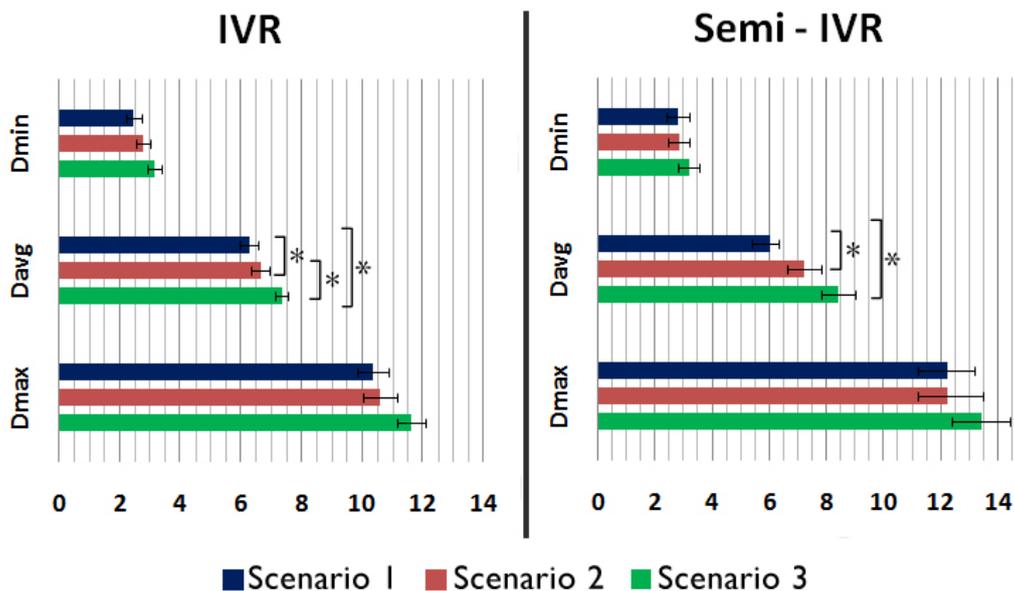


Figure 12: Minimum (D_{min}), Average (D_{avg}) and Maximum Distance (D_{max}) between the participant and child in each scenario. Error bars present standard error of means. * = $p < 0.05$.

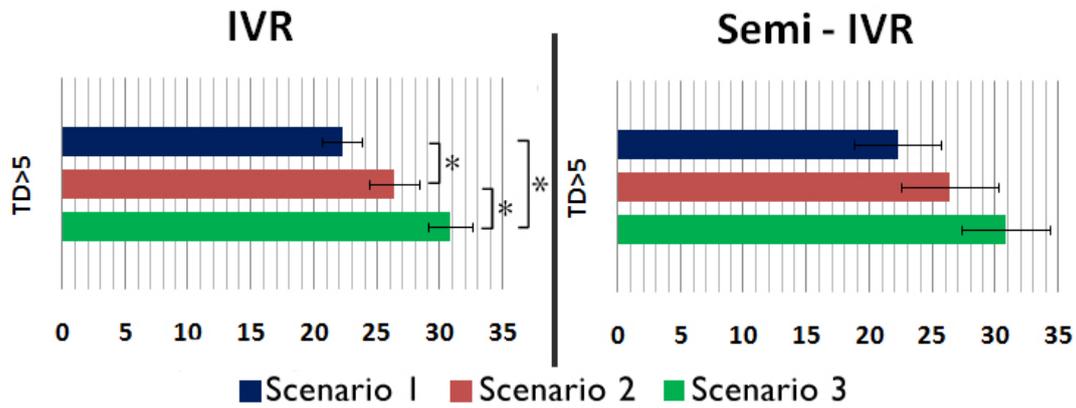


Figure 13: Mean time ($T_{D>5}$) that the participant remained more than five meters away from child in each scenario. Error bars present standard error of means. * = $p < 0.05$.

Data were tested for normality using the one-sample Kolmogorov-Smirnov test and the Shapiro-Wilk test. The results of these tests showed that almost all p-values were above 0.05, revealing that our datasets are normally distributed. Thus, for the statistical analysis of these data we used parametric tests for repeated-measures experiments' data.

To examine whether the four variables statistically differed between scenarios we conducted four repeated measures ANOVA with Greenhouse-Geisser correction. The variables D_{min} and D_{max} in both the IVR and the semi-IVR system had no statistically significant differences. Moreover, the $T_{D>5}$ in the semi-IVR also had no statistically significant differences. On the other hand, the variable D_{avg} did differ statistically significantly between scenarios in both systems (IVR: $p < 0.001$; semi-IVR: $p < 0.001$). Additionally, the variable $T_{D>5}$ in the IVR system did differ statistically significantly between scenarios ($p < 0.001$).

Post hoc tests using the Bonferroni correction revealed that for the IVR system, comparing S1 to S2, there was an increase in D_{avg} between the user and the child (6.27 ± 0.30 m. 6.65 ± 0.31 m. respectively) which was statistically significant ($p = 0.016$). D_{avg} in S3 was higher than in the other two scenarios (7.39 ± 0.22 m) which was statistically significant compared to S1 ($p = 0.001$) and S2 ($p = 0.017$). Studying the semi-IVR system using the Bonferroni correction, we found that there was also an increase in D_{avg} between the user and the child (6.00 ± 0.32 m, 7.20 ± 0.59 m, respectively) when comparing S1 with S2, which was not statistically significant ($p = 0.107$). However, D_{avg} in S3 was higher than in the other two scenarios (8.41 ± 0.60 m) which was statistically significant compared to S1 ($p = 0.001$), but not statistically significant compared with S2 ($p = 0.129$) (see supplemental material Tables 14-24).

Overall, the behavioral results suggested that participants kept the longest distance on average and spent the most time being far away from the child in S3, and kept the smallest distance and spent the least time being far away in S1. This effect exists in both IVR than semi-IVR but stronger in IVR. In other words, if being closer at all times to the child were our measurement of participants' performance, we have observed the pattern of $S1_Actual_Performance > S2_Actual_Performance > S3_Actual_Performance$ (i.e., participants performed the best in S1 and worst in S3).

5. DISCUSSION

The results of reliability and the validity tests together with the analysis of the validation question show that the statistical data are valid and reliable. The statistical results of the

studied factors yielded several important insights regarding the user interaction with a virtual crowd, which are similar for both systems (IVR and semi-IVR).

First of all, our study identified a strong pattern that when the crowd became more lifelike in their level of interactivity (collision avoidance, basic social behavior), the participants perceived a higher level of realism and reported a high level of presence. When collision avoidance between virtual characters and the user was enabled (S2), the user judged the characters, the environment and the whole VR system as more realistic and lifelike. Moreover, adding basic level of social interaction (S3) further contribute to this effect. This is consistent with the literature [16], [17] on human-virtual human interaction. This pattern is reflected in participants' perceived realism of the child, the crowd, and their reported level of presence, in both IVR and semi-IVR system. Note that the realism of the child in fact remained unchanged and therefore the perceived increased realism was indeed illusionary. With the other two measurements (*Aware_Others* and *Realism_Environment*), in our semi-IVR the same pattern was observed ($S3_Aware_Others > S2_Aware_Others > S1_Aware_Others$ and $S3_Realism_Environment > S2_Realism_Environment > S1_Realism_Environment$), but not in IVR. Given that, similar to *Child_Realism*, these two factors (crowd aware of each other, and the environment) were actually unchanged across conditions, this result suggested that in semi-IVR there was a stronger illusion of improved realism for the crowd and the environment. Such illusion is absent for IVR. We think it could be because in our IVR system participants were navigating with a wand rather than their body (walk-in-place) and therefore they were more focused on the target of their task, the child, rather than the crowd or the environment. This is further supported by the fact that their illusions of having increased realism of the child ($S3_Child_Realism > S2_Child_Realism > S1_Child_Realism$) were statistically significant in the IVR but not in semi-IVR system.

Secondly, when it comes to users' subjective evaluation of their performance, the results revealed an interesting pattern yet highly statistical significant: while the scenario with the highest level of interaction (S3, with both collision avoidance *and* social interaction) brought the highest level of comfort and easiness in both IVR and semi-IVR, enabling collision avoidance on its own (S2) produced a reduced feeling of comfort and easiness in navigating. This result suggests whilst we have a positive correlation between perceived realism and level of interactivity ($S3_Realism > S2_Realism > S1_Realism$), when it comes to perceived task performance there is a different pattern ($S3_Performance > S1_Performance > S2_Performance$). This suggested that when the crowd displays some low level interactively (collision but not social), it could have a negative impact on the whole experience. We believe that when collision avoidance was turned on, participants automatically (and subconsciously) raised their expectation in terms of social interaction with the crowd (see the literature where participants attributes mental states to virtual humans [3]). When such expected social interaction is non-exist, it hinders their perceived performance.

Moreover, when it comes to more objective measurements (users' behavior analysis), a surprising pattern was observed: we found a significant increase in the distance between the user and the child in the virtual world in the IVR system and the same tendency in the semi-IVR system when we enabled collision avoidance between the virtual characters and the participant. The growth in the distance was even bigger and statistically significant in both systems when we enabled both collision avoidance and interaction with the user. The observational analysis showed that participants were trying to "follow the rules" of the presented scenario in the VR environment. When the collision avoidance was enabled between the virtual characters and the participants, the latter were trying to avoid also the collisions in a much higher rate than when the collision avoidance was disabled. Moreover, when some virtual characters waved or verbally saluted the participants, they waved back or returned the

salutation. This may justify that both the interaction and the collision avoidance may reduce the user's performance regarding his/her primary goal, which was navigating into the VR environment with a certain target. Further, this is consistent with findings in [23].

Our study suggested that with both IVR and semi-IVR systems, collision avoidance should be accompanied with basic interaction between the user and the virtual crowd, such as verbal salutations, gaze, waving and other gestures. This may increase both the realism of the system and the feeling of comfort in the VR system, thereby enhancing the sense of presence of the user.

Our main aim was to investigate the relationship between human-crowd interactivity and presence and performance in VR. Here we present two different studies - one in IVR navigating via a wand and one in semi-IVR with walk-in-place. One potential drawback of our work is: although in this paper we discuss and compare the results from the two different systems, we understand that as they were two separate studies and therefore the results between these two systems should not be generalized as differences between IVR and semi-IVR systems. However, because of the interesting findings in this work, in future we would like to run more studies with controlled navigating method and number of participants to compare the two systems more systematically.

Further work is also necessary to study more factors that affect the user experience when they participate in populated IVE with virtual crowds. For example, it would be useful to examine the plausibility of different group formations and sizes in different types of environments in IVR systems. The environments could be open areas (such as an open-space mall or a square) or closed constrained areas (such as halls or corridors). The crowd might consist of different sizes of groups. The virtual humans could be individuals (singles), groups of two virtual pairs, groups of three (triples) or more. These two characteristics can be combined and studied using experiments, and by having participants evaluate each case. The outcomes can help to decide a plausible way for populating an IVR application of a specific type. The density of the virtual crowd can be a factor to be examined, since it can make a difference to the user's behavior. E.g. a dense crowd could incommode the navigation in the VE, affecting the participant's performance and sense of presence. Moreover, the extent of the interaction between the virtual characters and the participant is a factor that would be interesting to be further examined.

We also plan to study the effect of "secondary" behaviors of virtual humans on the immersed user's experience and sense of presence in an IVR application. These could be checking their watch, talking to each other, talking to their mobile phones etc. The results could be of great utility in the population of IVR applications with virtual crowds.

ACKNOWLEDGEMENTS

This research was partially funded by the Cyprus Research Promotion Foundation and the European Structural Funds for the "VR CAVE" project under contract IPE/NEKYP/0311/02.

REFERENCES

- [1] X. Pan, M. Gillies, C. Barker, D. M. Clark, and M. Slater, "Socially anxious and confident men interact with a forward virtual woman: an experimental study.," *PLoS One*, vol. 7, no. 4, p. e32931, Jan. 2012.
- [2] D. Pertaub, M. Slater, and C. Barker, "An experiment on public speaking anxiety in response to three different types of virtual audience," *Presence Teleoperators Virtual Environments*, vol. 11, no. 1, pp. 68–78, 2002.
- [3] D. Freeman, M. Slater, P. E. Bebbington, P. A. Garety, E. Kuipers, D. Fowler, A. Met, C. M. Read, J. Jordan, and V. Vinayagamoorthy, "Can virtual reality be used to investigate persecutory ideation?," *The Journal of Nervous and Mental Disease*, vol. 191, no. 8, pp. 509–514, 2003.
- [4] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments.," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 364, no. 1535, pp. 3549–57, Dec. 2009.
- [5] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis, "Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments," *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 6, pp. 583–598, 2001.
- [6] L. M. Wilcox, R. S. Allison, S. Elfassy, and C. Grelik, "Personal space in virtual reality," *ACM Transactions on Applied Perception*, vol. 3, no. 4, pp. 412–428, 2006.
- [7] J. Llobera, B. Spanlang, G. Ruffini, and M. Slater, "Proxemics with Multiple Dynamic Characters in an Immersive Virtual Environment," *ACM Transactions on Applied Perception*, vol. 8, no. 1, pp. 1–12, 2010.
- [8] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis, "Interpersonal distance in immersive virtual environments.," *Personality and Social Psychology Bulletin*, vol. 29, no. 7, pp. 819–833, 2003.
- [9] J. Bruneau, A.-H. Olivier, J. Pettré, J. Pettre, and J. Bruneau, "Going through, going around: a study on individual avoidance of groups," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 4, pp. 520–528, 2015.
- [10] C. Peters and C. Ennis, "Modeling groups of plausible virtual pedestrians.," *IEEE Computer Graphics and Applications*, vol. 29, no. 4, pp. 54–63, 2009.
- [11] R. McDonnell, C. Ennis, S. Dobbyn, and C. O'Sullivan, "Talking bodies: Sensitivity to desynchronization of conversations," *ACM Transactions on Applied Perception*, vol. 6, no. 4, pp. 1–8, 2009.
- [12] C. Ennis, R. McDonnell, and C. O'Sullivan, "Seeing is believing: body motion dominates in multisensory conversations," *ACM Transactions on Graphics*, 2010.
- [13] S. Huerre, J. Lee, M. Lin, and C. O'Sullivan, "Simulating believable crowd and group behaviors," *ACM SIGGRAPH ASIA 2010 Courses*, pp. 1–92, 2010.
- [14] C. Ennis, C. Peters, and C. O'Sullivan, "Perceptual effects of scene context and viewpoint for virtual pedestrian crowds," *ACM Transactions on Applied Perception*, vol. 8, no. 2, p. 10, 2011.
- [15] F. Friedmann, T. Schubert, and H. Regenbrecht, "The Experience of Presence: Factor Analytic Insights," *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 3, pp. 266–281, 2001.
- [16] M. Garau, D. Friedman, H. Ritter Widenfeld, A. Antley, A. Brogni, and M. Slater, "Temporal and Spatial Variations in Presence: Qualitative Analysis of Interviews from an Experiment on Breaks in Presence," *Presence Teleoperators Virtual Environments*,

- vol. 17, no. 3, pp. 293–309, 2008.
- [17] M. Slater, B. Lotto, M. M. Arnold, and M. V Sanchez-vives, “How we experience immersive virtual environments : the concept of presence and its measurement *,” *Anuario de Psicología*, vol. 40, no. 2773, pp. 193–210, 2009.
 - [18] M. Slater, C. Guger, G. Edlinger, R. Leeb, G. Pfurtscheller, A. Antley, M. Garau, A. Brogni, and D. Friedman, “Analysis of Physiological Responses to a Social Situation in an Immersive Virtual Environment,” *Presence Teleoperators Virtual Environments*, vol. 15, no. 5, pp. 553–569, Oct. 2006.
 - [19] N. Pelechano and C. Stocker, “Being a part of the crowd: towards validating VR crowds using presence,” *In Proceedings of the 7th international joint conference on Autonomous agents and mulitagent systems*, vol. 1, no. Aamas, pp. 12–16, 2008.
 - [20] J. Ahn, N. Wang, D. Thalmann, and R. Boulic, “Within-Crowd Immersive Evaluation of Collision Avoidance Behaviors,” in *VRCAI*, 2012, pp. 231–238.
 - [21] D. Thalmann and S. R. Musse, “Crowd Simulation,” *Springer*, 2007.
 - [22] D. Thalmann, H. Grillon, J. Maim, and B. Yersin, “Challenges in Crowd Simulation,” *2009 International conference CyberWorlds*, pp. 1–12, 2009.
 - [23] F. A. Sanz and L. Anatole, “Virtual Proxemics : Locomotion in the Presence of Obstacles in Large Immersive Projection Environments.” *2015 IEEE Virtual Reality (VR). IEEE, 2015.*
 - [24] M. Kyriakou, X. Pan, and Y. Chrysanthou, “Interaction with virtual agents - Comparison of the participants’ experience between an IVR and a semi-IVR system,” *2015 IEEE Virtual Reality (VR). IEEE, 2015.*
 - [25] D. W. Martin, *Doing psychology experiments*. Cengage Learning, 2008.
 - [26] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, “The CAVE: audio visual experience automatic virtual environment,” *Communications of the ACM*, vol. 35, no. 6. pp. 64–72, 1992.
 - [27] J. Smisek, M. Jancosek, and T. Pajdla, “3D with Kinect,” *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pp. 1154–1160, 2011.
 - [28] B. G. Witmer and M. J. Singer, “Measuring presence in virtual environments: A presence questionnaire,” *Presence Teleoperators virtual Environments*, vol. 7, no. 3, pp. 225–240, 1998.
 - [29] W. A. IJsselsteijn, “Presence: concept, determinants, and measurement,” *Proceedings of SPIE*, vol. 31, no. 0, pp. 520–529, 2000.
 - [30] K. Stanney, G. Salvendy, J. Deisinger, P. DiZio, S. Ellis, J. Ellison, G. Fogleman, J. Gallimore, M. Singer, L. Hettinger, R. Kennedy, J. Lackner, B. Lawson, J. Maida, A. Mead, M. Mon-Williams, D. Newman, T. Piantanida, L. Reeves, O. Riedel, T. Stoffregen, J. Wann, R. Welch, J. Wilson, and B. Witmer, “Aftereffects and sense of presence in virtual environments: Formulation of a research and development agenda,” *International Journal of Human-Computer Interaction*, vol. 10, no. 2, pp. 135–187, 1998.
 - [31] R. S. Kalawsky, S. T. Bee, and S. P. Nee, “Human Factors Evaluation Techniques to Aid Understanding of Virtual Interfaces,” *BT Technology Journal*, no. December, pp. 1–17, 1998.
 - [32] D. W. Schloerb, “A quantitative measure of telepresence,” *Presence Teleoperators Virtual Environ.*, vol. 4, no. 1, pp. 64–80, 1995.