Behavioral Measures of Copresence in Co-located Mixed Reality

Pierrick Uro, Florent Berthaut, Thomas Pietrzak, and Marcelo Wanderley

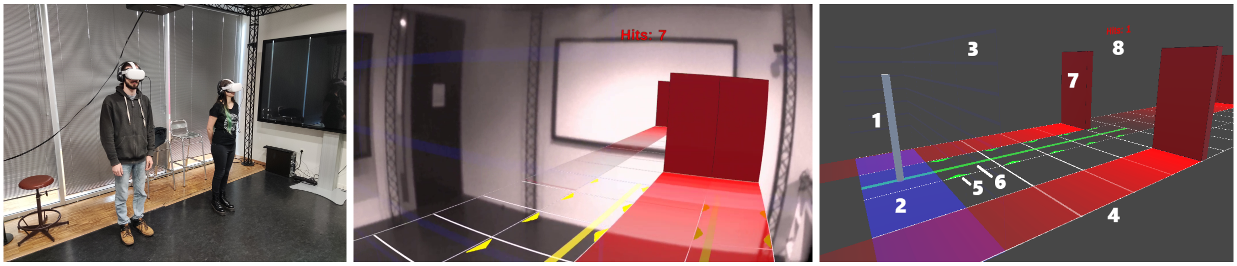

Fig. 1: Left: View of the physical space with two participants. Center: Point of view of a participant seeing several walls approaching them. Right: An annotated view of the virtual environment: The participant (1), task area (2), guard rails (3), lanes (4), lane feedback (5), position feedback (6), walls (7), and score (8)

Abstract—When several people are co-located and immersed in a mixed reality environment, they may feel like they share the virtual environment or not. This feeling of copresence, along with its parent dimensions of social presence and presence, has been mostly studied by relying on subjective measures gathered through questionnaires. As a way to address the drawbacks of this approach, we introduce a protocol to gather behavioral measures in the context of co-located mixed reality. As a pair of participants avoid obstacles moving towards them, their errors, gaze, interpersonal distance, and timing are measured. By combining subjective measures gathered through a questionnaire drawing from previous studies on social presence with behavioral measures, we demonstrate new ways to assess how users experience copresence. We illustrate this protocol by evaluating the effect of visual feedback on collaborators' activity. The results of this experiment suggest the capability of our protocol by revealing the effect of visual feedback on both objective and subjective measures.

Index Terms—Co-located mixed reality, copresence, shared virtual environment, collaboration.

1 Introduction

Co-located Mixed Reality (CMR) allows multiple users to share both physical and virtual environments. Such an opportunity has been made use of in various fields such as design [31], visualization [7, 35], or performance [47, 50]. However, co-located people may not always share the same virtual elements in various ways, like for instance people working on the same content but with their own instances and transformations of it [43, 55], or even using completely independent applications. Such cases raise a yet unanswered question: to what extent do users in CMR feel like co-located others are sharing the same experience?

This ambiguity is typically critical in Mixed Reality (MR) theatrical performances where the audience members perceive both physical performers and a virtual stage [29, 30, 48]. Because of the diversity and complexity of performers' actions, their variety of equipment (MR headsets, motion tracking), and the responses of the virtual environment, it can be difficult for the audience to perceive whether the performer is sharing their virtual environment or not. This difficulty can degrade their understanding of the performer's actions [14], and their feeling of sharing the virtual space, which may in turn negatively impact their engagement in and enjoyment of the performance [19, 33].

The psychological phenomenon of *presence* is one of the main aspects of evaluating these experiences, and it has been extensively studied and defined in contexts beyond the strict scope of MR. Specifically, we focus on *copresence*, defined as "*The degree to which the observer believes he/she is not alone and secluded, their level of peripherally or focally awareness of the other, and their sense of the degree to which the other is peripherally or focally aware of them.*" [11]. We do not conflate it with *social presence*, its parent dimension, which past studies have mainly employed to investigate whether a virtual entity is perceived like a person [39]. This inquiry becomes hollow when the other in question can be directly perceived as a physically present human as in CMR. The difference in focus brought by a co-located context makes this distinction key, since it also means that measuring tools and methods introduced for social presence will not be blindly translatable. Adapting known tools also needs to account for the context of MR, which has been pointed out to introduce new challenges when compared to Virtual Reality (VR) [45, 59, 68]. Notably, MR presents the potential for various elements in the environment to be felt with varying degrees of presence, creating a form of mixed presence [21]. Thus, we aim to determine how copresence applies to CMR, and how it can be measured in this context.

Subjective measures such as self-report questionnaires are omnipresent in the field of presence research. According to a recent review by Souza *et al.*, 86% out of 239 user studies on presence employed subjective measures while only 12% completed them with objective measures [61]. This is also the case for studies on social presence in MR [5, 16, 32, 71]. The issues of questionnaires are well known and often commented upon, pointing out their complexity with naive participants, that they can either only happen after the experience or risk breaking the sense of presence of the participant [28]. Another criticism is that a questionnaire is capable of eliciting in a participant the feeling that it was trying to measure in the first place [58].

The alternative is then to add complementary measures to the protocol, such as physiological or behavioral measures, each with their own advantages and disadvantages [28, 59]. Active tasks have been

- *Pierrick Uro is with the Université de Lille and McGill University. E-mail: pierrick.uro@univ-lille.fr*
- *Florent Berthaut and Thomas Pietrzak are with the Université de Lille. E-mail: [florent.berthaut|thomas.pietrzak]@univ-lille.fr*
- *Marcelo Wanderley is with McGill University. E-mail: marcelo.wanderley@mcgill.ca*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

found to affect physiological data through engagement in the task, the stress from questionnaires, and artifacts caused by movements [25]. However, while behavioral analysis allows for an approach that can be both qualitative and quantitative, it is heavily dependent on the context or task and must thus often be adapted [21], which is the case for CMR. Still, its versatility and non-intrusiveness make it a prime candidate for investigating copresence in CMR. It does however bring the challenge of selecting a task that is both able to provide behavioral measures of copresence, and to leave enough room for this feeling to fluctuate. The task and the design of the study also need to account for the likely issues of fatigue caused by wearing a headset and a learning effect known to happen in behavioral studies [3, 62].

We present a new protocol to allow the evaluation of copresence in CMR. It is built around the assumption that a higher sense of copresence will induce changes in behavior in users during a collaborative task. We hypothesize that users will perform physical movements to help one another avoid virtual obstacles coming towards them in CMR. Positional data is analyzed in conjunction with answers from the participants to interviews and a questionnaire drawing from established copresence questionnaires and adapted to the context of CMR. We apply this protocol in a user study comparing two conditions eliciting low and high copresence through positional visual feedback to identify which measures are relevant for its future iterations.

2 Related Work

Measurement techniques for presence and copresence can be classified into subjective and objective measures. In presence research, subjective methods are the most employed ones [61], mainly represented by questionnaires. Secondly, objective measures consist of either physiological or behavioral analysis. This part presents copresence and its parent dimensions, and then questionnaires and behavioral measures from the literature and how they can fit into the context of CMR.

2.1 Presence, Social Presence, and Copresence

Presence is a staple in the evaluation of virtual experiences and has been extensively studied and defined in contexts beyond the strict scope of MR. A way to unify these varying definitions is to see presence as two dimensions: *spatial presence* and *social presence* [21]. On the one hand, spatial presence, corresponds to the feeling that the user may have of being located within a specific environment [34]. It is commonly described as the feeling of "*being there*" [8, 49, 51], and is individual, leaving no room for the consideration of an other person.

On the other hand, social presence refers to the feeling of sharing a space with other people [39]. Investigations around this concept mostly focus on how a person determines whether an entity is perceived as *real* or not. In the context of CMR however, no doubt can remain on whether a physically present person is an actual person, and the question becomes instead whether they are perceiving the same environment. Moreover, a review including perception and cognition studies of presence between 2013 and 2018 reported that most of them consisted of remote collaborative systems [6]. Hence, studying this context must rely on more appropriate tools than those designed exclusively for spatial or social presence.

Still, the question fits as a sub-dimension of social presence, which has been further divided into *copresence*, *psychological involvement*, and *behavioral engagement* [10], which can be seen as three levels of social presence [9]. On a perceptual level, copresence is defined as the sensory awareness of an embodied other [38], and can even be extended to include a sense of mutual awareness [26]. On a subjective level, psychological involvement consists of how accessible the other seems, and in particular how aware one can be of the other's attentional engagement, emotional state, and comprehension [9]. And on an intersubjective level, behavioral engagement relates to someone's belief that their actions are interdependent with the ones of another. While copresence remains the root of our research question, other dimensions of this definition of social presence appear in part relevant to it, namely attentional and behavioral engagement, as well as perceived comprehension.

2.2 Presence and Copresence Questionnaires

Questionnaires attempting to evaluate presence are plenty and varied [28, 57]. The four most commonly found in the litterature are Witmer and Singer's Presence Questionnaire (PQ) [70], the Slater-Usoh-Steed Presence Questionnaire (SUS) [65], the MEC Spatial Presence Questionnaire (MEC-SPQ) [67], and the IGroup Presence Questionnaire (IPQ) [52]. Yet none of these questionnaires were explicitly designed for MR, which leads to researchers adding custom single-item questions to address aspects specific to MR [20]. This necessary approach coupled with the already large variety of questionnaires [57] furthers the difficulty in comparing publication results [20]. A questionnaire built specifically for MR is the Mixed Reality Experience Questionnaire (MREQ) [44], which probes the perceived relationships among the user, and the virtual or real objects, environment, or agents. Like the previously described questionnaires scoped around presence, the MREQ does not cover the nuances that make up social presence. Namely, it inquires about the existence of the various elements, and whether they felt as belonging together, but falls short on notions of behavioral interdependence or perceived comprehension with other agents.

These considerations are part of Biocca *et al.*'s definition of social presence, which served for the design of a dedicated questionnaire, the Networked Minds Measure of Social Presence (NMMSP) [11]. It has notably been applied for MR with a subset of its questions to focus on specific dimensions [27]. The NMMSP was adapted to the context of gaming, where communication is not the focus, in order to create the Social Presence in Gaming Questionnaire (SPGQ) with the notable addition of a mutual intention understanding question [18]. A similar approach can also be seen in order to fit presence questionnaires to the context of shared environments [15]. In our study of copresence in CMR, we draw from these two approaches, trimming and adjusting questions from the NMMSP in order to fit within the intended context.

However, the overwhelming use of questionnaires is often described and criticized [21, 24, 28, 57, 61]. Post-experience questionnaires can easily fall victim to limitations such as recall errors [21, 59, 66], intrusion [57], forcing the participants into a meta-view [58, 59], the influence of prior VR immersion [23], or the complexity of the concept either relying on the participant's interpretation or outright imposing the researcher's conceptual framework on the participant [58, 59]. Participants also need a frame of reference to answer Likert scale-based questionnaires, which is an issue with between-subjects experiment designs [46]. While relying on questionnaires presented during the experience may alleviate some of these limitations [13, 53], completing it with an objective measure such as physiological or behavioral analysis is recommended [21, 28, 59]. In this paper, we therefore investigate a behavioral approach that can be used in addition to questionnaires adapted for copresence in the context of CMR applications.

2.3 Behavioral Measures of Presence and Copresence

Much like Botvinick and Cohen's rubber hand illusion task for measuring immersion [12], many behavioral measures rely on an often stressful stimulus to elicit a reaction which will be measured. For instance, participants have been made to evolve close to a virtual pitfall, which allows to record instances of behaviors, like unprompted comments or where a reaction is incoherent, like stepping onto the pit [25]. In this scenario, more quantitative data can also be gathered through gait metrics such as stride length, width, and speed, along with the physiological measures of heart rate and galvanic skin response [41]. A similar scenario relies on flight phobia as a stimulus, which then allows measures through a subjective questionnaire, physiological data, and a behavior analysis [63]. This analysis is accomplished by a condition-blind observer, who reviews recordings and scores the speed, behavior, naturalness and reliability of the participant's walk before and after the scenario on a three point Likert scale. A unified scale was proposed for this class of scenarios, the Behavioral Presence Test in Threatening Virtual Environments (BPTT) [36]. It also relies on an external observer assessing the behavior of participants before and after they are exposed to the threat, like a fire or a pit. The scale consists of a list of anticipated behaviors, such as change in breathing pattern, or looking at the source of danger. While behavioral measures have also been used in MR by assessing risky behavior [36], all of these approaches are limited by their reliance on stressful stimuli, which constrains the possible tasks and whose applicability to copresence has not been verified.

There are also possible measures that do not rely on stressful situations [57], such as socially conditioned behaviors [56], pointing to an ambiguous object [60], comeback rate [64], or postural response [22]. However, none of these measures, which focus on individual actions or reactions, can be directly applied to a context of copresence. We therefore propose to establish behavioral measures around an indirectly collaborative task.

3 Protocol

In this paper, we propose a novel protocol for quantifying copresence in the context of CMR using behavioral measures.

Our approach draws inspiration from social behavior in everyday life situations. It assumes that the feeling of sharing a space with someone will induce changes in behavior. More specifically, it recreates the situation where, when walking alongside another person and seeing that this person will encounter an obstacle in the physical environment, one will tend to step aside in order to leave room for the other. This behavior depends both on the perception of the other person, but also on the perception of a shared physical environment that affects the other person.

Here we transpose this situation to a CMR environment for two users, in which the obstacles are virtual but participants remains physically present. We hypothesize that physical movements performed to help one another avoid virtual obstacles will reflect the level of copresence, since it will indicate that participants feel that they are together within the same physical and virtual spaces. If the copresence is lower, we expect participants to forget about each other and collectively fail to adapt their behavior and avoid obstacles.

Fig. 2: The proposed protocol recreates the physical situation where one person adapts their behavior to help another avoid an obstacle. We expect different behaviors depending on if the person is not aware of the other ($A \rightarrow B$) or if they are ($C \rightarrow D$)

In order to evaluate the effectiveness of this protocol, we conducted an experiment in which we used conditions that are known to have an effect on copresence, namely added visual feedback that act as an avatar to reinforce the integration of physical users in the virtual environment [17, 38].

3.1 Task

The task environment consists of four virtual lanes with a width of one meter each, and an area at the edge of these lanes with a depth of one meter (see Fig. 1). This is the task area, colored in blue, and is where the participants are instructed to stay. The task consists of avoiding a sequence of walls appearing on the other edge of the lanes every four seconds and advancing toward the participants at a speed of 2 m/s. A transparent grid surrounding the area gains opacity when the participant approaches it so they do not exit it by inadvertence, serving as guard rails, and a message indicating to go back to the blue area is displayed to the participants whenever they exit it. Walls will appear on the lanes as red rectangular obstacles and need to be dodged.

If a participant is on the same lane as an obstacle when a wall hits the blue area, it is recorded as a hit. Similarly, if the two participants are on the same lane when a wall reaches the blue area, a hit is added whether there was an obstacle in the lane or not. The goal of the participants is to minimize the amount of hits by the end of the experiment. Finally, they are asked not to cross each other during the task for safety and analysis reasons.

Wall Configurations There are six different kinds of walls configurations, as two obstacles are placed among four lanes. Each of these configurations only leaves two spaces for the participant out of the four lanes. With this setup, a wall is equivalent to a pair of positions which the participants must reach. Then, two successive walls represent a movement required of the participants, with a starting wall and an ending wall. In order to cover all possible movements, each unique pair of wall configurations must be represented. It is however possible to reduce the amount of total walls necessary by considering that an ending wall can serve as a starting wall at the same time. We can thus build five sequences of walls that each covers all movements possible with 31 walls, including a first wall that serves only as a starting wall. Each pair experiences all five sequences in counter-balanced order.

Forcing Walls A particularity of the task is the existence of *forcing* situations, which is whenever a participant must move in order to leave room for their partner so they can avoid a hit (see Figure 3). Such situations require the participant to not only focus on their own movement but to also take into account the existence of the other person within the virtual environment. These situations are precisely the ones where the tendency to leave room to the other can be measured, and are the ones where measures and analysis should be focused. Flooding the pair of participants with *forcing* situations is however risking that its resolution by them becomes too conscious. This is why each wall sequence consists of all 30 possible situations instead of a repetition of the 10 *forcing* ones.

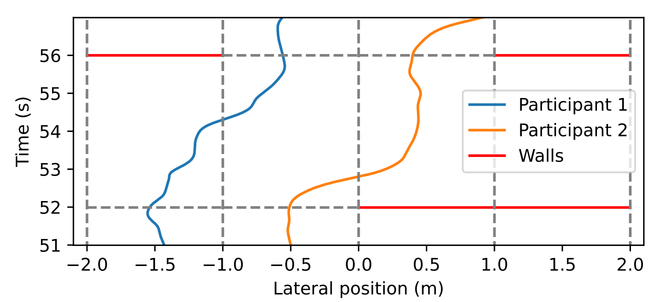

Fig. 3: Position of two participants going through a forcing wall. They are positioned by the first wall at t=52, then the right participant could stay in the same lane without hitting a wall at t=56, but must step right in order to leave room to their neighbor.

Within or Between? Within-subjects design has been shown to detect existing embodiment effects that between-subjects design was not able to reveal [46], making it the favored design for this kind of protocol. It would allow less individual variance in the data, reduce the amount of participants necessary, and enable richer discussions during a post-experience interview. It does however lengthen experimental sessions, which should be avoided in contexts prone to cybersickness such as MR, and is vulnerable to transfer across conditions. This task is especially liable to learning effects across conditions, as participants may identify *forcing* situations over time and develop strategies, which makes within-subjects design inappropriate. Many other behavioral studies of presence or social presence in MR have gone with a between-subject design presumably because of this learning effect [3, 37, 62]. In order to alleviate this limitation, an extra *perspective block* is added after the participants fill in their questionnaires. Its purpose is to allow them to experience alternative conditions in order to enrich the discussion that follows in the semi-structured interview, and it should not be included in the quantitative analysis.

3.2 Metrics

The protocol integrates both subjective and objective measures through the use of a questionnaire, an interview, and behavioral measures.

3.2.1 Questionnaire

The post-task questionnaire provides a subjective measure of copresence by drawing mainly from the NMMSP questionnaire [11]. Since the NMMSP was designed in direct relation to the definitions of social presence and copresence used for this study, it serves as a perfect basis for a questionnaire focused on copresence in CMR. Since the task is a game, it is logical to also consider de Kort *et al.*'s adaptation of it to gaming situations [18], in particular its behavioral engagement section, which is most relevant to copresence.

The self-report questionnaire consists of 24 questions (see Section 6) built out of four dimensions presented in random order:

Cognitive Level The separation between three Cognitive Levels Situation, Behavior, and Intention mirrors de Kort *et al.*'s addition of a question around mutual intention understanding to the NMMSP. It also matches the Psychological Involvement in part described by Biocca *et al.* as the belief of having insight into the intentions, motivations, and thoughts of the other [11].

Dependence Level Decomposing Dependence Level into Awareness and Dependence allows to include Awareness as a key part of the definition of copresence by Biocca *et al.* [11], that also relates to Attentional Allocation and Mutual Awareness, two factors indicating copresence. On the other hand, Dependence relates to the Behavioral Engagement dimension of social presence, with its factors being Behavioral Interdependence, Mutual Assistance, and Dependent Action.

Target The Target of the question can be either the participant themselves (Reported), or their partner (Attributed). This follows most copresence and social presence questionnaires which feature mirrored versions of their questions, one where the participant reports on their own feelings, and one where they conjecture on the feelings of the others.

Positive/Negative A redundancy is added by presenting both a positive and negative form of each question. In the analysis, the negative answer is inverted, and any pair of answer that suggests an inconsistency with a difference of more than two points is dismissed.

3.2.2 Interview

The interview is semi-structured in order to nurture discussion and gather comments and opinions from the participants themselves that can complete the interpretation of results. It consist of different parts aiming at various aspects of the experience of the participants.

The first questions are aimed directly at the condition that is being evaluated. Their answers should help supporting analytical results with the own opinions and feelings of the participants themselves.

Then, in case specific behaviors or errors committed by the participants were identified by either the interviewer or participants, they are pointed out and discussed. Namely, asking for the perceived cause of specific errors helps to understand the cognitive processes of the participants by having them confront their subjective and objective comprehension of their neighbor [14]. This serves as a stepping stone for a more general discussion on the potential strategies that both participants developed. Encouraging the participants to exchange on these strategies is helpful for exposing potential mismatches in their respective approaches of the task.

A few questions give an opportunity to the participants for expressing to what extent they felt copresence in relation with their partner. Examples of such questions are to ask them if they felt like they were playing by themselves, or if they stopped feeling that a person was standing next to them.

Finally, potential anticipated limitations are addressed, like by inquiring on the interface or the effort necessary for the task.

While some of these answers could be collected through a questionnaire, obtaining them through semi-open questions enables the emergence of issues or effects not anticipated by the experimenters.

3.2.3 Performance

A link between task performance has been pointed out by a meta-analysis of 80 studies that validated a model in which social presence positively impacts *flow*, a mindstate of high focus [69], which in turn positively impacts task performance [40]. It shows that task performance is also subject to task effort, which is impacted by trust, and in turn social presence. We can thus conjecture a potential link between copresence, a sub-component of social presence, and task performance.

In this protocol, participants are given an explicit goal, which is to avoid the walls by not being in the same lane as their partner or the wall when it reaches the task area. A performance measure is thus obtained by using the ratio of walls on which an error was committed. Since both participants share a score, and an error may be caused by either or both of them, the performance metric is also shared among the pair.

3.2.4 Trajectory

The position over time of the participants can be modeled as a trajectory in order to allow analysis. Here, a trajectory is defined as the sequence of lateral directions a participant moves towards. The transitions between these directions follow a hysteresis model. By default, a participant is still, and when their velocity reaches a threshold of 0.5 m/s, they are considered to be moving in that direction. Once they are moving, they need to go under a velocity of 0.25 m/s in order to be considered still again. The trajectory is further simplified by allowing a change in direction only 0.5 s after the last one. This model was established through trial and error until each block was adequately characterized while displaying stable trajectories based on visual confirmation, an example of it can be seen in 4. With it, the trajectory of each participant can be analyzed through its temporality and the distance kept between both.

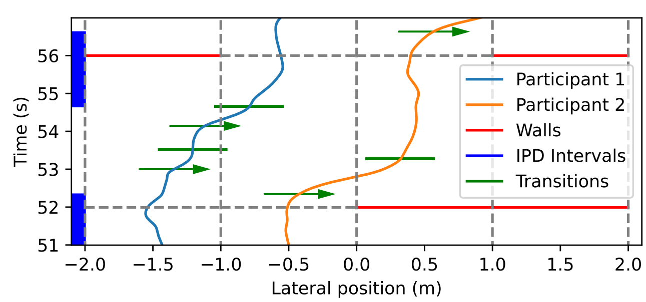

Fig. 4: Position of two participants going through a forcing wall (the same as Figure 3). Trajectory transitions are represented as green lines or arrows, with a line representing transitioning to a still state, and an arrow being in movement towards the direction it points to. The blue rectangles to the left indicate the intervals on which the interpersonal distance is computed. Note that the short interval around t=54 is not computed as participants are not in adjacent lanes, and the interval lasts for less than one second.

Temporality Gait characteristics such as stride length, width, and speed have been used as behavioral measures in threatening environments [36,41], and while there is little threat in this scenario, a link may be found between movement speed and copresence. A higher speed can also be interpreted as a form of task performance, or even a marker of flow, both previously linked to positive social presence [40].

For each trial two delays are computed around the starting and ending wall (see subsection 3.1). The Starting Delay is the time spent between when the starting wall reaches the task area and the moment the participant begins moving. If the participant started moving before the wall, the Starting Delay will be negative. Complementarily, the Ending Delay is the time spent between the end of the participant's movement and the arrival of the ending wall. Again, if the participant stops moving after the wall reaches the task area, the Ending Delay will be negative. Walls featuring an error or with an unexpected trajectory are excluded from the analysis as outliers.

Interpersonal distance The maintaining of an interpersonal distance can be seen as a mark of a greater sense of copresence or social presence [2, 4]. Notably, distance between collaborating people has been shown to be different with co-located and remote participants despite a similar self-report of copresence [42].

A measure of interpersonal distance can be extracted by averaging the distance between the participants during all periods where they are both considered still and in adjacent lanes. Such periods that lasted less than one second are excluded so that both participants are considered to be in resting position.

3.2.5 Gaze

Studies of gaze in social situations in virtual environments point towards an avoidance of another's gaze with higher social presence [1]. For a different reason, in the context of this protocol, looking towards the other may also be interpreted as a lack of mutual awareness. The reasoning behind this interpretation is that looking towards the other in a game where the rules dictate their position can be seen as a mark of the participant not being fully confident or aware of the potential position of their partner. In both cases, more time spent towards the other is expected to reveal a lower sense of copresence.

To measure the gaze of a participant, the direction towards which their headset faced is recorded during the entire task. We define 6 areas towards which this direction can point (see Figure 5): FORWARD-CENTER (FC), FORWARD-DOWN (FD), FORWARD-UP (FU), PARTNER (P), OPPOSITE-PARTNER (OP), and BACK (B). The FORWARD area (F) spans 120° horizontally. This value is based on the 60° a participant must turn before a person standing on a perpendicular line next to them enters their field of view, which gives an indication that the participant is looking far enough to be able to see their partner. This area is further divided into FC, spanning 100°, with FU, and FD, respectively above and below. The choice of the size of the division is also made with the field of view in mind, as a participant must look 50° downwards before their feet enter their vision. This distinction allows to separate when a participant is looking at the walls ahead from when they are looking at their own feet to adjust their position. The B area also spans 120° and is located directly opposite F. The direction of the lateral areas, P and OP, depend on the side of the participant during the task. For instance if the participant is on the left of their partner, P is towards the right.

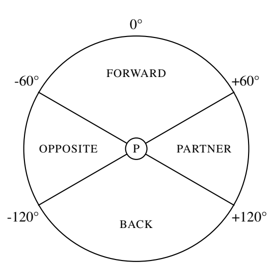

Fig. 5: The delimitation of the gaze areas. The Forward area is vertically split between Center, Up, and Down

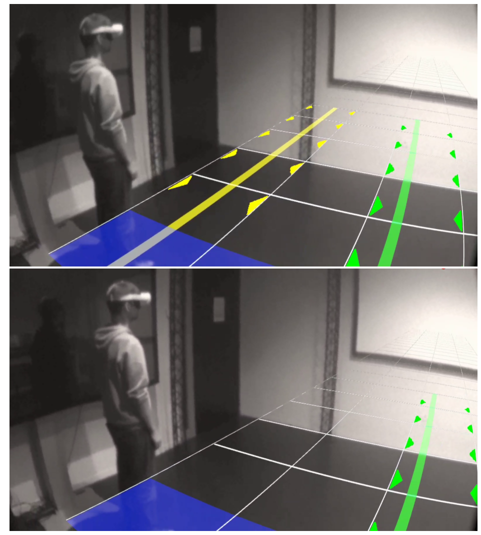

Fig. 6: Both feedback conditions (FEEDBACK top, NO FEEDBACK bottom), a participant can see their own feedback in both conditions

4 Experiment: Effect of Visual Feedback

In order to evaluate the effectiveness of our protocol, we implement a first study on the effect of visual feedback that inform of the others' activity. We design this experiment on the basis that the existence of a visual feedback of the other person causes a higher level of copresence [17]. Having a condition with high copresence and one with lower copresence allows the evaluation of each metric selected in our protocol.

On the lanes, discreet and continuous feedback inform the participants of their position in relation to the virtual environment. Arrows are displayed inside of the lane in which a participant is and move whenever they go into a different lane, providing discreet feedback. Simultaneously, a line extending from the blue area follows the lateral position of the participant, serving as continuous feedback. The two conditions relate to these feedback modalities: in the condition FEEDBACK, each participant is able to see their own position feedback as well as their partner's, while in the condition NO FEEDBACK, the participants can not see their partner's feedback and can only see their own.

With these two conditions, the experiment follows a between-subjects design with 2 Feedback × 5 blocks × 30 Walls (of which 10 are *forcing*).

4.1 Apparatus

Both participants wear a Meta Quest 2 headset which communicate with a server hosted on a laptop and each other over OSC on a shared Wifi network. The headsets allow MR by providing access to the black and white external camera feed, which is displayed under the virtual environment.

The participants also wear Srhythm NC25 headphones that are playing pink noise throughout the task. The virtual environment was developed on Unity (2020.3.36f1) with the Oculus package for MR and the extOSC package for communication between the headsets and server.

The physical room is smaller than the lanes on which the obstacles advance, meaning that they appear as if behind a physical wall. Orthogonal subdivisions were added to provide positional context for the virtual obstacles in order to limit this issue.

4.2 Procedure

After the participants are equipped with the headsets and headphones, the virtual environment is described to them, along with the task and its rules. The participants are instructed not to communicate verbally with each other during the task in order to allow potentially conflicting approaches among the pair. For this intent, white noise is played on the headphones worn by the participants in order to mask the sound of the footsteps of their partner and most attempts to communicate with sounds. They are asked to adjust the volume of the noise themselves, so it is loud enough that it covers footsteps as demonstrated by the experimenter, but not so loud that it becomes uncomfortable.

A short training phase then begins, where they avoid ten walls each blocking one lane at a time. After an optional break, the participants go through all five blocks also interleaved with optional breaks during which the participants may take off their headsets and headphones if they feel the need to. Each block consists of 31 walls, each covering any two of the four lanes, and lasts 135 s, for a total of 12 min including the training and without the pauses. Once all blocks are completed, the participants take off their headsets and headphones, and each one completes an individual questionnaire. An additional block using the condition that was not selected is then completed in order to grant to the participants a perspective of the other condition. Finally, an interview between the experimenter and both participants is conducted.

4.3 Participants

In total, 15 pairs of participants were recruited. One pair was however excluded from behavioral analysis because they exhibited outlier behavior by playfully going into the lane of their partner without necessity. The data from a total of 28 participants was therefore analyzed, of which 7 female, and 21 male, aged between 19 and 68 with a median of 30 (std=12.3). Seven pairs completed the task in the NO FEEDBACK condition, and seven in the FEEDBACK condition. Of these pairs, 3 were with mixed gender, leaving 2 all female pairs, and 9 all male. Gender was balanced between conditions, with 4 female participants with feedback, and 3 without.

4.4 Results

For the following analysis, only data that occured within a *forcing* situation are considered. After a Shapiro-Wilk normality test, we ran two-way mixed ANOVA with FEEDBACK as a between-subject factor with two levels, and BLOCK as a within-subject factor with five levels. Where necessary, a Greenhouse-Geisser sphericity correction was applied. Post-hoc analysis was ran with Bonferroni correction. If a lack of normality could be found, we conducted instead a non-parametric Kruskal-Wallis test on the same factors. Results are reported as statistically significant when $p < 0.05$.

4.4.1 Performance

The error rate was computed for each pair in *forcing* situations.

With a Shapiro-Wilk value of 0.692 ($p < 0.001$), a significant difference from normal distribution was found, preventing the use of ANOVA. Instead, a Kruskal-Wallis test showed that performance was significantly affected by both FEEDBACK ($H(1) = 5.675$, $p < 0.017$) and BLOCK ($H(4) = 21.462$, $p = 0.001$), as seen in Figure 7.

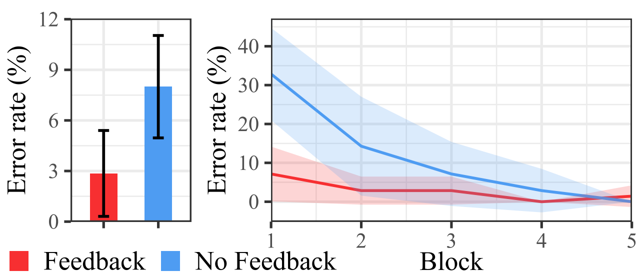

Fig. 7: Error rates of both conditions. Error bars and shaded areas are 95% confidence intervals.

4.4.2 Trajectories

Trajectories were computed using the model and values described in Part 3.2.4.

Temporality The Shapiro-Wilk test was not significant for either Starting or Ending Delays ($W = 0.981$, $p = 0.191$ and $W = 0.221$, $p = 0.622$ resp.). A main effect of BLOCK can be found both on Starting ($F_{2.6,67.8} = 8.50$, $p < 0.001$) and Ending Delays ($F_{2.9,75.0} = 9.50$, $p < 0.001$) (see Figure 8). No effect was found for FEEDBACK on Starting ($F_{1,26} = 0.26$, $p = 0.618$) or Ending Delays ($F_{1,26} = 0.05$, $p = 0.819$). There was also no interaction betwen FEEDBACK and BLOCK for either (resp. $F_{2.6,67.8} = 0.78$, $p = 0.493$ and $F_{2.9,75.0} = 0.72$, $p = 0.540$).

For Starting Delay, a post-hoc analysis shows a significant difference between the first block and blocks 3, 4, and 5 (resp. $p = 0.002$, $p < 0.001$, $p < 0.001$), and between the second block and blocks 4 and 5 (resp. $p = 0.041$ and $p = 0.022$).

For Ending Delay, a post-hoc analysis shows a significant difference between the first block and each other block ($p = 0.003$ for block 2, and $p < 0.001$ for every other).

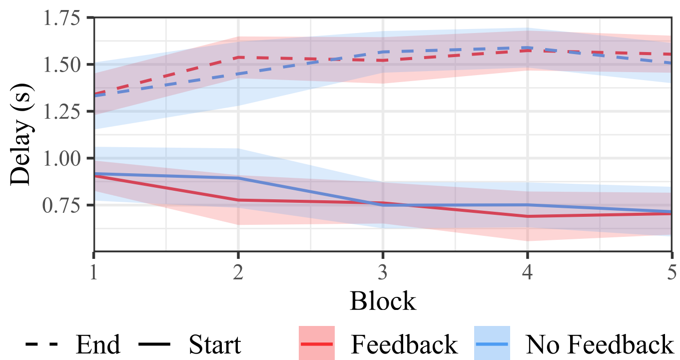

Fig. 8: Starting and Ending Delay of both conditions for each block with standard error. Shaded areas are 95% confidence intervals.

Interpersonal Distance The normality test was passed ($W = 0.953$, $p = 0.610$). No main effect was found for either FEEDBACK ($F_{1,12} = 1.56$, $p < 0.236$), BLOCK ($F_{2.7,31.5} = 2.48$, $p < 0.086$), or their interaction ($F_{2.7,31.5} = 1.11$, $p < 0.355$).

4.4.3 Gaze

The time spent looking towards each area during the 10 *forcing* walls each block was gathered, totaling 40 seconds per block. A pair was excluded for this metric as their recording failed on the last three walls, compromising their total time spent during forcing walls.

Focusing on the FORWARD-CENTER and PARTNER areas, a Shapiro-Wilk test showed an absence of normal distribution ($W = 0.884$, $p < 0.001$ and $W = 0.836$, $p < 0.001$ resp.), leading to a Kruskal-Wallis test for both areas on factors FEEDBACK and BLOCK. In the FORWARD-CENTER area, a significant effect was found for FEEDBACK ($H(1) = 35.679$, $p < 0.001$), and none for BLOCK ($H(4) = 0.658$, $p = 0.956$). In the PARTNER area too, a significant effect was found for FEEDBACK ($H(1) = 51.501$, $p < 0.001$), and none for BLOCK ($H(4) = 0.280$, $p = 0.991$).

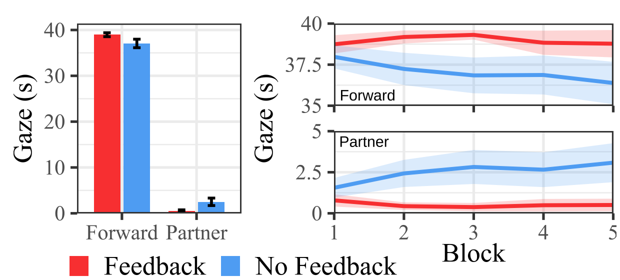

Fig. 9: Time spent facing towards the Partner and Forward areas in both conditions for forcing walls. The total duration of these trials is 40 s. Error bars and shaded areas are 95% confidence intervals.

4.4.4 Questionnaire

The mean global score in the FEEDBACK condition is 2.15 (std=0.71), and 1.85 (std=0.57) in the NO FEEDBACK condition (see Figure 10). A pairwise t-test comparison shows a statistically significant effect of feedback on the global score (s=3.077, p=0.002). A more detailed analysis of the dimensions in the questionnaire, that can be seen in Figure 11, suggests that there are more positive answers and a quasi-total absence of negative answers in the FEEDBACK condition.

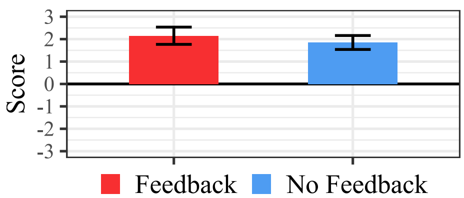

Fig. 10: Global Questionnaire Score for both conditions. Error bars are 95% confidence intervals.

4.4.5 Correlation between questionnaire and explicit errors

We computed the correlation between the errors rate and the global questionnaire score, which results in a moderate negative correlation (Pearson's r=-0.415, p=0.028), *i.e.*, the lower the error rate, the higher the self-reported copresence.

5 DISCUSSION

In this section, we discuss the effectiveness and the limitations of behavioral measurements for quantifying copresence in CMR. We provide insights on the importance of territories and learning, and we reflect on how to implement our protocol in other contexts.

5.1 Behavioral measures of Copresence in CMR

Our goal with this protocol was to be able to measure copresence in CMR through behavioral analysis. The role of this first study is to validate the protocol by selecting a condition whose effect on copresence is expected, to confirm this effect through a questionnaire, and to evaluate the behavioral measures prepared for it. This first study demonstrates the effectiveness of behavioral measures and our approach, through the following evidence.

First, we observe that the questionnaire was able to show a difference in the global score between both conditions, with the FEEDBACK condition scoring better than NO FEEDBACK. This means that participants felt an overall stronger mutual awareness and interdependence with their partners when their positional information was made available to them within the virtual environment. This can be understood through statements from participants in the FEEDBACK condition stating "*The marker became more you than you*", or "*After a while, I considered the line to be an extension of my partner's position*". On the other hand, a participant in the NO FEEDBACK condition commented that they needed to "*leave the game to look at the other before entering it again*", confirming that the lack of feedback deteriorated the sense of sharing the environment, while its presence kept the other inside the virtual environment as an elementary unidimensional avatar. These results suggest that the feeling of copresence was increased with the addition of visual feedback, which was the base hypothesis for testing the protocol. At the same time, our measures show an effect of this visual feedback on the behavioral errors. Error rates were higher in the absence of visual feedback, confirming our hypothesis that a reduced representation of the other users within the virtual environment would impact the overall performance in avoiding walls.

Finally, for both areas of interest in gaze, FORWARD-CENTER and PARTNER, a main effect of FEEDBACK was found, along with an interaction with BLOCK. Participants without feedback spent more time looking toward their partner and less time looking forward. They reported during the interview that looking towards their partner served to check for potential conflict, and that while some were able to guess where the other was, they would still look either for reassurance or entertainment.

This combination of results can be interpreted as follows: given a stronger sense of copresence (induced through added feedback), participants pay more attention to each other's behavior and adapt their trajectories to collectively avoid walls. This is confirmed through a statistically significant correlation between the errors from our behavioral measures and the global score from our shared experience questionnaire, which tends to reinforce the link between the participants' behavior and their feeling of copresence. One could argue that the added visual feedback alone might explain the effect that we observe. However the fact that participants in the NO FEEDBACK group were able to perfectly complete the task at the end of block 5 leads us to the interpretation that differences in performance are more related to thinking or remembering that one has to step aside, due to the reminder of the presence of the other participant. This also aligns with the testimony of the participants.

However, we note that no effect of FEEDBACK was found for either delays or interpersonal distance. This can mean that neither of these measures was adequate for this task, or that the amount of data is too small to find an effect. What might cause the lack of results from these behavioral measures for this task is that it relies heavily on the participants being highly aware of their position and timing. This contrasts with a study finding an effect of co-location on mean and minimum clearance, whose task involved movement without making it such a key part of the goal since it involves the gathering of objects in the environment [42]. This suggests that either the avoidance task or the trajectory metrics should be adapted in order to better fit each other.

5.2 Errors and Copresence

This study is also already able to give insight on how the feeling of copresence can fluctuate. During the post-task interviews, participants were asked about the reasons for their errors and close calls. While some evoked fatigue, or a clear disagreement where the other was considered but their intention was unexpected, many errors stem from a false feeling of relief. A participant commented "*As far as I was concerned, I was through, but I didn't think about my partner.*", a reoccurring attitude which the participants sometimes had enough time to correct without triggering a hit. This betrays occasional lapses in copresence, which are revealed through conflict-inducing *forcing* walls. Another said "*I see a free lane in front of me so I figure I'm in the clear for four walls.*", which reveals a tendency by participants to look ahead and anticipate their movements, sometimes without considering their partner.

This supports a vision of a continuous copresence fluctuating throughout the task in a similar way to the "Breaks in Presence" (BIP) approach of measuring presence [59]. However in this case, instead of a high sense of presence dropping because of an event, a low sense of copresence can surge up. More specifically, the sense of copresence slowly degrades with fatigue or intense focus on the walls, and jumps back up when the other is brought back into consideration, either because of an error or a physical contact with the other, as has sometimes happened during trials. The appearance of visual feedback on the other person in this case serves as a near-constant reminder of the other's presence, which limits the degradation of copresence over time.

5.3 Learning effect

Our results show a strong learning effect across blocks, which needs to be taken into account for future development of the protocol. A learning effect was detected through BLOCK for errors, starting and ending delays. Over time, participants committed less errors, started and stopped moving earlier. This is especially visible for NO FEEDBACK groups, which start at a higher error rate than FEEDBACK ones, as can be seen in Figure 7. This effect could be because of the participants learning the task and developing a strategy, or because they build a stronger sense of the other over time allowing them to move without hesitation.

The creation of a strategy is commonly reported by participants during the interview. Some explain that they identified that the external lane to their side would always be theirs and they would always stay

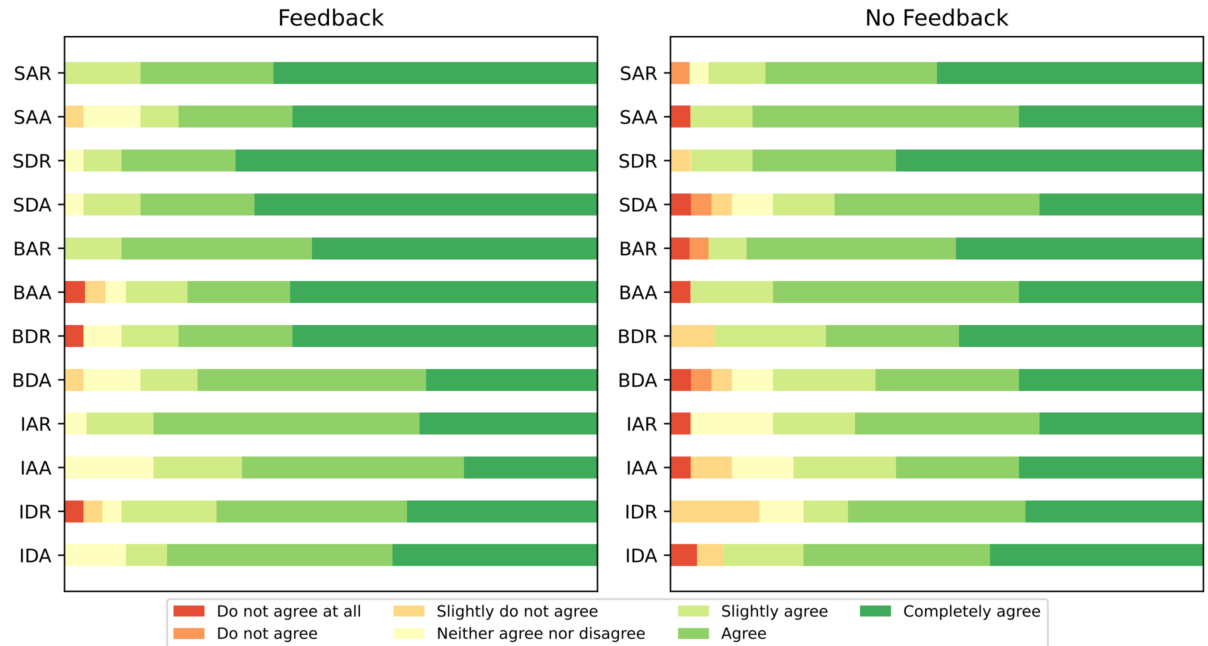

Fig. 11: Answer proportion for each question. Each letter represents a value from the three dimensions **S**ituation/**B**ehavior/**I**ntention, **A**wareness/**D**ependence, and **A**ttributed/**R**eported (see Section 3.2.1). For instance, **BAR** refers to **B**ehavior **A**wareness **R**eported.

there when it was open, hinting at the emergence of territories [54]. Others solved the puzzle by noticing that with only two available lanes and no right of crossing each other, there was only one option left, which was to always take the leftmost lane if they were on the left. Finally, a few improved this approach by considering time. They identified that there were cases where they had to move quickly to allow room for their neighbor as soon as possible, and cases where they should move slowly so their partner has time to step aside.

It also seems like the sense of sharing an environment can be learned. A participant noted that "*At some point you realize that you play as two, then you don't need to think*". Another noticed that their understanding of their partner's behavior was stable, but that they could guess their intention as they learned the game.

This calls for an iteration of the protocol with too much complexity for a strategy to be defined during the task, for instance by allowing the participants to cross each other, increasing the number of lanes, or restricting them to certain participants. The observation of a cross-condition learning effect afforded by the perspective blocks also substantiates the decision of a between-subject design for the study. Future applications of the task should also consider this potential learning effect in their analysis.

Another important aspect involved in the learning process is the gaze. A statistically significant interaction between FEEDBACK and BLOCK for gaze and a post-hoc analysis reveal a form of learning for where participants in the NO FEEDBACK condition look during the task. The sharp increase in PARTNER gazing matching the sharp decrease in error rate between the first and second blocks in the NO FEEDBACK condition hints towards a correlation between the two. However some participants reported during the interview that they did not understand before the second block that they could look directly at their partner, which can partially explain the difference in gazing. Other participants also reported a lack of understanding of the rules and did not realize that they should not stand in the same lane until their first error. These two possibilities were less likely to impact participants in the FEEDBACK condition, as the position of their partner was available in the the virtual environment without having to turn their head. The feedback of both participants turning red when in the same lane could also have helped some understand the rules before committing an error.

5.4 Other contexts and copresence factors

The proposed behavioral measure can also be transposed to evaluate copresence in application contexts which do not rely on full body movements. In order to do so, the key components that need to be preserved are: 1) spatial territories that participants can appropriate and share; 2) a collaborative task that may provoke conflicts within these territories; 3) sufficient independence between participants within that task so that their levels of copresence can vary.

For example, in a context of mid-air manipulations, one could envision a line of virtual boxes between two participants facing each other and in which they have to reach from above to touch one target each as fast as possible. Moving the two targets between boxes can then be used to test if participants will leave the closest target to one another to ensure success. In the case of a low level of copresence, as in the wall dodging task, we expect participants to forget about the other and fail to leave room for them.

Following the same principle, other contexts could also be envisioned, covering selection, manipulation and navigation within virtual content, and with more than two users.

Within this context or others, more potential factors that could be studied include environment asymmetry, inspired by actual events encountered during the experiment. We observed a participant stepping forward, out of the task area by 4 meters, without either them or their partner noticing because they were both too focused on the task. Leveraging the focus required by the avoidance task with the spacial asymmetry would thus reveal its potential effect on copresence in CMR. By exploiting the continuous nature of behavioral measures, it is possible to progressively or abruptly displace the virtual environments in relation to the physical area. By controlling the triggering of potential Breaks in Co-presence (BIC), as an equivalent of Breaks In Presence (BIP) [59] for shared environments, the behavior pre- and post-BIC could be compared, which would give a measure for the impact of the displacement causing the BIC. It will thus be possible to investigate the tolerance to spatial misalignment in shared CMR environments.

A different kind of asymmetry relying on the same ideas of BIC is one where participants punctually encounter different walls. This is based on the observation of a participant seeing their partner not taking any action at all to avoid a wall in their lane and immediately thinking that the system encountered an issue. The participants noticing the differences by themselves, as well as changes in their behavior could serve as measures of their awareness of their neighbor and understanding of their behavior for other factors.

5.5 Limitations

Our contribution suffers from limitations which should be addressed in future versions of the protocol. First the FEEDBACK and NO FEEDBACK conditions are still entangled with the measures of task performance (which could be improved by the visual feedback alone) and gaze (influenced by the visual feedback). Our protocol therefore requires additional interviews and questionnaires to remove any doubt on what is exactly measured. This issue could be addressed by increasing the complexity of the environment, *e.g.* with more lanes and options, in order to enable a greater diversity of solutions, therefore reducing the learning effect and the influence of visual feedback.

Second, trajectories did not provide strong evidence of an effect of the two conditions. Other designs for the obstacles and movements, together with alternative methods for analyzing these trajectories, might help reinforce their usefulness in the estimation of the feeling of a shared mixed-reality experience.

Third, we recognize that the number of data points may seem limited. It is due to the attempt of mitigating the learning effect in the experiment by using a between-groups design. Here too, a more complex task with multiple solutions might reduce the learning effect and allow for within-subject design with more participants.

Finally, the role of the passthrough quality on the feeling of copresence can be argued. The impact of photographic realism on copresence is unclear [39], unlike behavioural realism, which would be similar for low or high quality camera passthrough or visual see-through. Still, this asks for further investigation.

6 CONCLUSION

In this paper we propose a novel experimental protocol to evaluate the feeling of copresence in a CMR experience through both objective, *i.e.* behavioral measures, and subjective, *i.e.* questionnaires, aspects. We design an experimental task in which users collectively interact with a virtual environment in a way that highlights their perception of others in the physical and virtual environment through positional visual feedback.

Our results suggest that our protocol provides a complementary way of measuring the experience of users in CMR by confirming the effect of feedback via a questionnaire and measuring effects on performance and gaze.

This protocol is meant to be easily iterable and expandable in order to accommodate the evaluation of varied factors. As discussed above, it can also be transposed to other application contexts. Further research with this protocol will aim to strengthen its validity with new factors such as interface asymmetry.

SUPPLEMENTAL MATERIALS

The positive versions of the questions in the questionnaire and a recording of a trial can be found as supplemental materials.

ETHICS STATEMENT

This study was approved by the authors' ethical review board (Ethics Committee in Behavioral Science of the Université de Lille) with the reference number 2023-743-S122.

REFERENCES

[1] 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*, 10(6):583–598, Dec. 2001. doi: 10.1162/105474601753272844 5

[2] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis. Interpersonal Distance in Immersive Virtual Environments. *Personality and Social Psychology Bulletin*, 29(7):819–833, July 2003. doi: 10.1177/0146167203029007002 5

[3] J. N. Bailenson, R. Guadagno, E. Aharoni, A. Dimov, A. Beall, and J. Blascovich. Comparing behavioral and self-report measures of embodied agents' social presence in immersive virtual environments. *Proceedings of the 7th Annual International Workshop on PRESENCE*, Jan. 2004. 2, 4

[4] J. N. Bailenson, K. Swinth, C. Hoyt, S. Persky, A. Dimov, and J. Blascovich. The Independent and Interactive Effects of Embodied-Agent Appearance and Behavior on Self-Report, Cognitive, and Behavioral Markers of Copresence in Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 14(4):379–393, Aug. 2005. doi: 10.1162/105474605774785235 5

[5] C. Basdogan, C.-H. Ho, M. A. Srinivasan, and M. Slater. An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction*, 7(4):443–460, Dec. 2000. doi: 10.1145/365058.365082 1

[6] R. A. J. d. Belen, H. Nguyen, D. Filonik, D. D. Favero, and T. Bednarz. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering*, 3(2):181–223, 2019. doi: 10.3934/ElectrEng.2019.2.181 2

[7] H. Benko, E. W. Ishak, and S. Feiner. Collaborative mixed reality visualization of an archaeological excavation. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 132–140. IEEE, 2004. 1

[8] F. Biocca. The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments [1]. *Journal of Computer-Mediated Communication*, 3(2):JCMC324, Sept. 1997. doi: 10.1111/j.1083-6101.1997.tb00070.x 2

[9] F. Biocca and C. Harms. Defining and measuring social presence: Contribution to the networked minds theory and measure. *Proceedings of the Fifth Annual International Workshop on Presence*, Jan. 2002. 2

[10] F. Biocca, C. Harms, and J. K. Burgoon. Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators and Virtual Environments*, 12(5):456–480, Oct. 2003. doi: 10.1162/105474603322761270 2

[11] F. Biocca, C. Harms, and J. L. Gregg. The Networked Minds Measure of Social Presence : Pilot Test of the Factor Structure and Concurrent Validity. 2001. 1, 2, 4

[12] M. Botvinick and J. Cohen. Rubber hands 'feel' touch that eyes see. *Nature*, 391(6669):756–756, Feb. 1998. doi: 10.1038/35784 2

[13] S. Bouchard, G. Robillard, J. St-Jacques, S. Dumoulin, M.-J. Patry, and P. Renaud. Reliability and validity of a single-item measure of presence in vr. In *The 3rd IEEE international workshop on haptic, audio and visual environments and their applications*, pp. 59–61. IEEE, 2004. 2

[14] O. Capra, F. Berthaut, and L. Grisoni. Have a SEAT on Stage : Restoring Trust with Spectator Experience Augmentation Techniques. In *Designing Interactive Systems (DIS)*. Eindhoven, Netherlands, July 2020. doi: 10.1145/3357236.3395492 1, 4

[15] J. Casanueva and E. Blake. The Effects of Group Collaboration on Presence in a Collaborative Virtual Environment. In *Virtual Environments 2000*, pp. 85–94. Springer Vienna, Vienna, 2000. doi: 10.1007/978-3-7091-6785-4_10 2

[16] T. Combe, R. Fribourg, L. Detto, and J.-M. Normand. Exploring the Influence of Virtual Avatar Heads in Mixed Reality on Social Presence, Performance and User Experience in Collaborative Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 30(5):2206–2216, May 2024. doi: 10.1109/TVCG.2024.3372051 1

[17] E. A. J. Croes, M. L. Antheunis, A. P. Schouten, and E. J. Krahmer. Teasing apart the effect of visibility and physical co-presence to examine the effect of CMC on interpersonal attraction. *Computers in Human Behavior*, 55:468–476, Feb. 2016. doi: 10.1016/j.chb.2015.09.037 3, 5

[18] Y. de Kort, W. IJsselsteijn, and K. Poels. Digital games as social presence technology : development of the social presence in gaming questionnaire (SPGQ): 10th Annual International Workshop on Presence (PRESENCE 2007), October 25-27, 2007, Barcelona, Spain. *Proceedings of the 10th Annual International Workshop on Presence, October 25-27, 2007*, pp. 195–203, 2007. 2, 4

[19] J. Decock, J. Van Looy, L. Bleumers, and P. Bekaert. The pleasure of being (there?): an explorative study into the effects of presence and identification on the enjoyment of an interactive theatrical performance using omnidirectional video. *AI & SOCIETY*, 29(4):449–459, Nov. 2014. doi: 10.1007/s00146-013-0487-6 1

[20] T. Drey, M. Rietzler, and E. Rukzio. Questionnaires and Qualitative Feedback Methods to Measure User Experience in Mixed Reality, Apr. 2021. 2

[21] W. M. Felton and R. E. Jackson. Presence: A Review. *International*

Journal of Human–Computer Interaction, 38(1):1–18, Jan. 2022. doi: 10.1080/10447318.2021.1921368 1, 2

[22] J. Freeman, S. E. Avons, R. Meddis, D. E. Pearson, and W. IJsselsteijn. Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. *Presence: Teleoperators & Virtual Environments*, 9(2):149–164, 2000. 3

[23] J. Freeman, S. E. Avons, D. E. Pearson, and W. A. IJsselsteijn. Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence*, 8(1):1–13, 1999. 2

[24] D. Friedman, A. Brogni, C. Guger, A. Antley, A. Steed, and M. Slater. Sharing and Analyzing Data from Presence Experiments. *Presence: Teleoperators and Virtual Environments*, 15(5):599–610, Oct. 2006. doi: 10.1162/pres.15.5.599 2

[25] M. Gandy, R. Catrambone, B. MacIntyre, C. Alvarez, E. Eiriksdottir, M. Hilimire, B. Davidson, and A. C. McLaughlin. Experiences with an AR evaluation test bed: Presence, performance, and physiological measurement. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 127–136. IEEE, Seoul, Korea (South), Oct. 2010. doi: 10.1109/ISMAR.2010.5643560 2

[26] E. Goffman. *The presentation of self in everyday life*. The presentation of self in everyday life. Doubleday, Oxford, England, 1959. 2

[27] J. G. Grandi, H. G. Debarba, and A. Maciel. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 127–135. IEEE, Osaka, Japan, Mar. 2019. doi: 10.1109/VR.2019.8798080 2

[28] S. Grassini and K. Laumann. Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review. *Frontiers in Psychology*, 11, 2020. 1, 2

[29] A. Henson. *We're In This Together: Embodied Interaction, Affect, and Design Methods in Asymmetric, Co-Located, Co-Present Mixed Reality*. thesis, Carnegie Mellon University, June 2019. doi: 10.1184/R1/8235539.v1 1

[30] B. Kates. Incandescent Edges of the Future: Performance Creation with Virtual, Augmented and Carbon Realities. Master's thesis, Arts, Sept. 2020. doi: 10.11575/PRISM/38180 1

[31] L. Kent, C. Snider, J. Gopsill, and B. Hicks. Mixed reality in design prototyping: A systematic review. *Design Studies*, 77:101046, 2021. 1

[32] K. Kim, R. Schubert, J. Hochreiter, G. Bruder, and G. Welch. Blowing in the wind: Increasing social presence with a virtual human via environmental airflow interaction in mixed reality. *Computers & Graphics*, 83:23–32, Oct. 2019. doi: 10.1016/j.cag.2019.06.006 1

[33] H.-G. Lee, S. Chung, and W.-H. Lee. Presence in virtual golf simulators: The effects of presence on perceived enjoyment, perceived value, and behavioral intention. *New Media & Society*, 15(6):930–946, Sept. 2013. doi: 10.1177/1461444812464033 1

[34] M. Lombard and M. T. Jones. Defining Presence. In M. Lombard, F. Biocca, J. Freeman, W. IJsselsteijn, and R. J. Schaevitz, eds., *Immersed in Media: Telepresence Theory, Measurement & Technology*, pp. 13–34. Springer International Publishing, Cham, 2015. doi: 10.1007/978-3-319-10190-3_2 2

[35] T. Mahmood, W. Fulmer, N. Mungoli, J. Huang, and A. Lu. Improving information sharing and collaborative analysis for remote geospatial visualization using mixed reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 236–247. IEEE, 2019. 1

[36] E. Malbos, R. M. Rapee, and M. Kavakli. Behavioral Presence Test in Threatening Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 21(3):268–280, Aug. 2012. doi: 10.1162/PRES_a_00112 2, 3, 5

[37] M. R. Miller and J. N. Bailenson. Social Presence Outside the Augmented Reality Field of View. *Frontiers in Virtual Reality*, 2:656473, Sept. 2021. doi: 10.3389/frvir.2021.656473 4

[38] K. L. Nowak and F. Biocca. The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 12(5):481–494, Oct. 2003. doi: 10.1162/105474603322761289 2, 3

[39] C. S. Oh, J. N. Bailenson, and G. F. Welch. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI*, 5, 2018. 1, 2, 9

[40] S. Parameswaran and R. Kishore. A Social Presence Model of Task Performance: A Meta-Analytic Structural Equation Model. *AMCIS 2017 Proceedings*, Aug. 2017. 4, 5

[41] L. Phillips, V. Interrante, M. Kaeding, B. Ries, and L. Anderson. Correlations Between Physiological Response, Gait, Personality, and Presence in Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 21(2):119–141, Apr. 2012. doi: 10.1162/PRES_a_00100 2, 5

[42] I. Podkosova and H. Kaufmann. Co-presence and proxemics in shared walkable virtual environments with mixed colocation. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11. ACM, Tokyo Japan, Nov. 2018. doi: 10.1145/3281505.3281523 5, 7

[43] T. Porcino, S. A. Ghaeinian, J. Franz, J. Malloch, and D. Reilly. Design of an Extended Reality Collaboration Architecture for Mixed Immersive and Multi-surface Interaction. In B. Göbl, E. van der Spek, J. Baalsrud Hauge, and R. McCall, eds., *Entertainment Computing – ICEC 2022*, pp. 112–122. Springer International Publishing, Cham, 2022. doi: 10.1007/978-3-031-20212-4_9 1

[44] H. Regenbrecht, C. Botella, R. Baños, and T. Schubert. Mixed Reality Experience Questionnaire (MREQ) -Reference. Feb. 2017. 2

[45] H. Regenbrecht and T. Schubert. Measuring Presence in Augmented Reality Environments: Design and a First Test of a Questionnaire, 2002. doi: 10.48550/arXiv.2103.02831 1

[46] G. Richard, T. Pietrzak, F. Argelaguet, A. Lécuyer, and G. Casiez. Within or between? comparing experimental designs for virtual embodiment studies. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 186–195. IEEE, 2022. 2, 3

[47] A. Rostami, C. Rossitto, and A. Waern. Frictional Realities: Enabling Immersion in Mixed-Reality Performances. pp. 15–27, June 2018. doi: 10.1145/3210825.3210827 1

[48] C. Salter, T. Thomasson, and P. Uro. Animate: A Theatrical Exploration of Climate Transformation through the Medium of Extended Reality (XR). In *28th International Symposium on Electronic Art*, pp. 642–651. Ecole des arts decoratifs - PSL, Paris, France, June 2024. doi: 10.69564/ISEA2023-83-full-Salter-et-al-Animate 1

[49] C. Sas, G. M. P. O'Hare, and R. Reilly. Presence and task performance: an approach in the light of cognitive style. *Cognition, Technology & Work*, 6(1):53–56, Feb. 2004. doi: 10.1007/s10111-003-0145-8 2

[50] R. Schlagowski, D. Nazarenko, Y. Can, K. Gupta, S. Mertes, M. Billinghurst, and E. André. Wish You Were Here: Mental and Physiological Effects of Remote Music Collaboration in Mixed Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–16. ACM, Hamburg Germany, Apr. 2023. doi: 10.1145/3544548.3581162 1

[51] R. Schroeder. *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. Springer Science & Business Media, Dec. 2001. 2

[52] T. Schubert, F. Friedmann, and H. Regenbrecht. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, June 2001. doi: 10.1162/105474601300343603 2

[53] V. Schwind, P. Knierim, N. Haas, and N. Henze. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–12, 2019. 2

[54] S. D. Scott, M. S. T. Carpendale, and K. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, CSCW '04, pp. 294–303. Association for Computing Machinery, New York, NY, USA, Nov. 2004. doi: 10.1145/1031607.1031655 8

[55] M. Sereno and T. Isenberg. Subjective Views in Co-Located Augmented Reality-Initial Design. Posters of IEEE Visualization, Oct. 2020. 1

[56] T. B. Sheridan et al. Musings on telepresence and virtual presence. *Presence Teleoperators Virtual Environ.*, 1(1):120–125, 1992. 3

[57] R. Skarbez, F. P. Brooks, Jr., and M. C. Whitton. A Survey of Presence and Related Concepts. *ACM Computing Surveys*, 50(6):1–39, Nov. 2018. doi: 10.1145/3134301 2, 3

[58] M. Slater. How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments. *Presence*, 13(4):484–493, Aug. 2004. doi: 10.1162/1054746041944849 1, 2

[59] M. Slater, D. Banakou, A. Beacco, J. Gallego, F. Macia-Varela, and R. Oliva. A Separate Reality: An Update on Place Illusion and Plausibility in Virtual Reality. *Frontiers in Virtual Reality*, 3, 2022. 1, 2, 7, 8

[60] M. Slater, M. Usoh, and Y. Chrysanthou. The influence of dynamic shadows on presence in immersive virtual environments. In *Virtual Environments' 95: Selected papers of the Eurographics Workshops in Barcelona,*

Spain, 1993, and Monte Carlo, Monaco, 1995, pp. 8–21. Springer, 1995. 3

[61] V. Souza, A. Maciel, L. Nedel, and R. Kopper. Measuring Presence in Virtual Environments: A Survey. *ACM Computing Surveys*, 54(8):163:1–163:37, Oct. 2021. doi: 10.1145/3466817 1, 2

[62] B. Spanlang, T. Fröhlich, V. F. Descalzo, A. Antley, and M. Slater. The Making of a Presence Experiment: Responses to Virtual Fire. 2007. 2, 4

[63] F. Steinicke, G. Bruder, K. Hinrichs, A. Steed, and A. L. Gerlach. Does a Gradual Transition to the Virtual World increase Presence? In *2009 IEEE Virtual Reality Conference*, pp. 203–210. IEEE, Lafayette, LA, Mar. 2009. doi: 10.1109/VR.2009.4811024 2

[64] S. Thie and J. van Wijk. A general theory on presence. *1st Int. Wkshp. on Presence*, 1998. 3

[65] M. Usoh, E. Catena, S. Arman, and M. Slater. Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments*, 9(5):497–503, Oct. 2000. doi: 10.1162/105474600566989 2

[66] J. Van Baren. Measuring presence: A guide to current measurement approaches. *Deliverable of the OmniPres project IST-2001-39237*, 2004. 2

[67] P. Vorderer, W. Wirth, F. R. Gouveia, F. Biocca, T. Saari, L. Jäncke, S. Böcking, H. Schramm, A. Gysbers, T. Hartmann, C. Klimmt, N. Ravaja, A. Sacau, T. Baumgartner, and P. Jäncke. MEC Spatial Presence Questionnaire (MEC-SPQ). 2004. 2

[68] I. Wagner, W. Broll, G. Jacucci, K. Kuutii, R. McCall, A. Morrison, D. Schmalstieg, and J.-J. Terrin. On the Role of Presence in Mixed Reality. *Presence: Teleoperators and Virtual Environments*, 18(4):249–276, Aug. 2009. doi: 10.1162/pres.18.4.249 1

[69] J. Webster, L. K. Trevino, and L. Ryan. The dimensionality and correlates of flow in human-computer interactions. *Computers in Human Behavior*, 9(4):411–426, Dec. 1993. doi: 10.1016/0747-5632(93)90032-N 4

[70] B. G. Witmer and M. J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, June 1998. doi: 10.1162/105474698565686 2

[71] F. Zaman, C. Anslow, A. Chalmers, and T. Rhee. MRMAC: Mixed Reality Multi-user Asymmetric Collaboration. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 591–600. IEEE, Sydney, Australia, Oct. 2023. doi: 10.1109/ISMAR59233.2023.00074 1