




# In case of doubt, one follows one's self: the implicit guidance of the embodied self-avatar

Loën Boban , Ronan Boulic , and Bruno Herbelin 

**Abstract**—The sense of embodiment in virtual reality (VR) is commonly understood as the subjective experience that one's physical body is substituted by a virtual counterpart, and is typically achieved when the avatar's body, seen from a first-person view, moves like one's physical body. Embodiment can also be experienced in other circumstances (e.g., in third-person view) or with imprecise or distorted visuo-motor coupling. It was moreover observed, in various cases of small or progressive temporal and spatial manipulations of avatars' movements, that participants may spontaneously follow the movement shown by the avatar. The present work investigates whether, in some specific contexts, participants would follow what their avatar does even when large movement discrepancies occur, thereby extending the scope of understanding of the self-avatar follower effect beyond subtle changes of motion or speed manipulations. We conducted an experimental study in which we introduced uncertainty about which movement to perform at specific times and analyzed participants' movements and subjective feedback after their avatar showed them an incorrect movement. Results show that, when in doubt, participants were influenced by their avatar's movements, leading them to perform that particular error twice more often than normal. Importantly, results of the embodiment score indicate that participants experienced a dissociation with their avatar at those times. Overall, these observations not only demonstrate the possibility of provoking situations in which participants follow the guidance of their avatar for large motor distortions, despite their awareness about the avatar movement disruption and on the possible influence it had on their choice, and, importantly, exemplify how the cognitive mechanism of embodiment is deeply rooted in the necessity of having a body.

**Index Terms**—Virtual Reality, virtual embodiment, sense of body ownership, sense of agency, self-avatar follower effect.

## 1 INTRODUCTION

In Virtual Reality (VR), one exists within the virtual environment and interacts with their surrounding through a virtual representation of themselves: their avatar. In such a context, one can experience the illusion that one's own body is substituted by the seen virtual body. The resulting strong relationship binding one to their avatar is characterized by the Sense of Embodiment, commonly acknowledged to arise from the combination of the feelings of owning and controlling the virtual body, and of being located inside it. In practice, the subjective feeling of embodying the virtual body is strongly experienced by participants immersed in virtual reality (e.g. with a Head Mounted Display) when they see the avatar's body from a first person perspective and observe that it is moving congruently with their physical body. In this context, it is considered that embodiment would be experienced by participants if some of these conditions are met: their avatar follows their movements, has the same appearance, or other perceptions expected from their physical body is also experienced through the avatar [6, 7, 13, 27, 30]. What is however less studied, although probably fundamental, is that the experience of embodiment seems to be more of a dialogue between the participants and their avatars, similarly to how we spontaneously experience our relation to our body [8]. The bidirectionality of the participant-avatar relationship is perhaps best illustrated by the participants' spontaneous reaction to the "loss" of their avatar when a movement distortion is artificially introduced (voluntarily or due to technical limitations). In such cases participants would, when possible, compensate for the observed discrepancy between their avatar and their physical body posture [4, 11] while otherwise, under certain conditions, would follow the avatar's movement [9, 25, 43]. When occurring, this phenomenon, called the *Self-Avatar Follower Effect* by Gonzalez-Franco et al. [25], results in a reduction of the multi-sensory discrepancies between one's physical body and the visual feedback,

which in turn contributes to the preservation of the sense of embodiment. Conversely, very large or sudden disruptions of the visuo-motor congruency between one's movements and the displayed virtual body might still trigger a *Self-Avatar Follower Effect* [25] but in that case can lead to a break in embodiment [31, 38, 42]. It is thus within a subtle range of postures and movement distortions that the spontaneous corrections characterizing the *self-avatar follower effect* seem to be triggered, and the challenge is to identify the requirements for this motor attraction to be elicited, regarding the nature of the movement distortion, the magnitude of the manipulation, and favorable context (e.g., type of task, previous movements). The *self-avatar follower effect* was previously investigated for movement distortions preserving a strong correlation between the user's initial movement and the avatar's one, such as when subtly distorting a reaching-forward movement [14, 25], or when manipulating only the speed of participant's movements [43] or the synchronicity of the avatar animation [9]. Concerning more drastic movement distortions, interesting behaviors were anecdotally observed in our lab in situations when participants would see a pre-animated full-body avatar viewed from a first-person perspective; without specific instructions about how to react while observing the virtual body co-localized with themselves but moving on it's own, some participants would explicitly mirror the seen actions. This also relates to the observation that distortions that can be noticed can still be tolerated by participants without impeding their subjective experience of embodiment [39, 40]. Said otherwise, participants could notice at first that their avatar is initiating a movement, and accept in a second time this guidance to maintain the embodiment. Conditions of uncertainty or doubt would thus be favorable as, without definitive motor planning, volitional control is not contradicted.

The present work aims at extending the scope of understanding of the self-avatar follower effect, beyond subtle changes of motion or speed manipulations, to include larger discrepancies, by exploring how VR users perceive or tolerate drastic movement distortions, such as when their avatar acts by itself and performs a movement independently from users' actual motion. To study this, we systematically induced situations in which participants were unsure about which movement to perform, while their avatar would execute a possible alternative, and analysed their behaviour.

• Loën Boban, Ronan Boulic, and Bruno Herbelin are with École Polytechnique Fédérale de Lausanne (EPFL). E-mail: {loen.boban | ronan.boulic | bruno.herbelin}@epfl.ch.

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.xxx/TVCG.201x.xxxxxx

## 2 RELATED WORK

When considering avatars in virtual reality, the *Sense of Embodiment* designates the subjective experience rising from the combination of the *Sense of Agency*, the *Sense of Body Ownership*, and the *Sense of Self-Location* [30]. In this framework, the *Sense of Agency* refers to the subjective experience of being in control of the avatar [26,47], while the *Sense of Body Ownership* designates the feeling of owning the virtual body [6, 8]. Finally, the *Sense of Self-Location* corresponds to the sensation of being located within the displayed avatar's body [18, 35]. From a multi-sensory perspective, the sense of embodiment toward a body  $B$  is interpreted as *the sense that emerges when  $B$ 's properties are processed as if they were the properties of one's own biological body* [30]. Consistently, as virtual embodiment is commonly understood as the result of substituting one's physical body with a virtual counterpart, it depends on the multi-sensory congruency between one's real body cues and the visual feedback of the displayed virtual body. Therefore, a strong sense of embodiment is usually achieved when one is provided with an avatar they can see from a first person perspective [15, 16], which overlays with one's own body, and that moves congruently with their real body [19].

Interestingly, once successfully elicited, virtual embodiment is robust to small multi-sensory incongruencies. For instance, small visuo-motor discrepancies between the avatar's movements with regard to one's actual movement would lead to multi-sensory disruptions that can be tolerated without breaking the sense of agency [20, 24, 29, 32, 41]. However, large discrepancies between one's movements and their avatar's movements would result in strong visuo-motor conflicts, which could cause a break in agency, in turn leading to a break in embodiment [31, 39, 40, 42]. Similarly, the sense of embodiment is robust to small visuo-proprioceptive mismatches but can be broken by large discrepancies. For example, in the case of a virtual hand illusion, the sense of ownership can be elicited toward a hand that does not overlay with the real hand, but the ownership illusion significantly drops when the virtual hand is placed in a very unnatural position [45].

Of particular interest, not only does the sense of embodiment benefit from the multi-sensory alignment between the real body cues and the virtual body ones but, in case of multi-sensory mismatches and under specific conditions, virtual embodiment can also be associated to unintentional movements diminishing the multi-sensory conflicts which, in turn, contributes to preserve it. For instance, in a Rubber Hand Illusion context [10], while inducing a visuo-proprioceptive mismatch by positioning the rubber hand at a distance from the hidden real hand, it was observed that participants experiencing the ownership illusion toward the rubber hand exhibited an automatic tendency to drift their physical hand in the direction of the rubber counterpart [2]. This finding was later replicated in virtual reality, with participants exerting forces towards the virtual hand when its position was offsetted compared to their real hand, and little to no forces when the virtual and the real hand were co-located [33].

Extending these findings from visuo-proprioceptive mismatch to visuo-motor discrepancies, similar behaviors were reported in virtual reality, showing the users tendency to follow the movements of their avatar, thereby minimizing potential mismatches between their real body cues and the visual feedback they are presented. In their work, Gonzalez-Franco et al. asked participants to reach forward with their right arm while their movement was projected on a predefined axis corresponding to a maximum angular distortion of 30°, thereby resulting in a rightward shift of the final hand's position [25]. In this context, it was observed that participants adjusted their movement in a way reducing the spatial offset between their real arm and the avatar's one. This behavior, called the *self-avatar follower effect* and described as "how when a spatial offset is introduced between the real and the virtual body, and the system allows for compensations, participants automatically act to reduce the spatial offset" [25], was later extended by considering a vertical distortion equivalent, further showing that participants tended to both horizontally and vertically move towards their avatar without having been instructed to do so [14]. Such motor behavior, that have

been associated to motor contagion and mimicry [9, 25], echoes previous observations made while participants were asked to draw straight lines while their virtual hand drew ellipses [12]. In such a context, participants were unaware that their actual motor performance was "attracted" toward the seen movement, suggesting that a conscious detection of the distortion can be associated with an unintentional motor response.

From a cognitive perspective, the *self-avatar follower effect* is believed to stem from the need to minimize conflicts caused by the alteration of self-perception consecutively to the embodiment of an avatar [25, 36]. This hypothesis draws from the active inference theory, which posits that actions contribute to the overall goal of reducing prediction errors [23, 44]. In the context of the *self-avatar follower effect*, as one cannot directly influence the visual feedback to mitigate the multisensory conflict, one may move their real body toward the avatar to reduce the discrepancy between their visually perceived body and their physical one.

Further extending these results to full-body movements, it was shown that when alternating between a real-time animation of the avatar and slow-motion, participants tended to quickly adjust their movement speed to align with the avatar's one, thereby maintaining a consistent visuo-motor experience [43]. This latter phenomenon was further investigated for asynchronous cyclic and repetitive movements by asking participants to perform repetitive upper- and lower-body movements while their avatar animation was either congruent or out-of-phase [9]. In this case, participants tended to synchronize with their avatar, while some reported that their avatar did not influence their movements, despite the behavioral effect.

Put together, these studies collectively demonstrate the embodied VR users' tendency to adapt their real movements to match with their virtual avatar's movements, whether it involves adjusting a trajectory, adapting to a given speed, mirroring actions, or syncing rhythm. These heterogeneous observations support the notion of a bidirectional motor relationship between users and their avatars in virtual reality, as they emphasize the influence of the avatar's movement on the user's motor behavior.

## 3 METHODS

To explore the possibility of inducing the *self-avatar follower effect* for large visuo-motor discrepancies, a specific task must be carefully designed. Indeed, no-one is expected to follow their avatar if the seen gesture patently prevents the completion of an ongoing task. For instance, if the instructions explicitly require participants to lift an arm in front of them, seeing the avatar acting by itself doing a very different movement, such as lifting a knee, would break the sense of embodiment and, likely, no follower-effect would be observed [9]. Therefore, to investigate whether the *self-avatar follower effect* could occur for substantial movement distortions, we experimentally induced uncertainty about which movement should be performed at given times, to put the participant in front of several plausible options. In such circumstances, if participants hesitate between different movements, the self-avatar could act as a guide and influence their choice in favor of the movement done by the avatar. One way to address this requirement is to push participants to the limits of their memory span for sequences of movements [46]. We thus designed an experimental task asking participants to memorize and repeat sequences of various movements, made to be particularly hard to recall. To counter the fact that people have different memorization abilities [3, 5, 46, 48] and ensure a uniform difficulty of our task, the sequence length was adjusted for each participant according to their individual memory performance.

### 3.1 Virtual environment

For the entire duration of the experiment, participants were immersed in a virtual environment by wearing a head mounted display (Valve Index) and body trackers (subsection 3.2). Instructions were given by a conversational agent acting as a virtual experimenter during the whole experiment. This agent was animated and communicated with

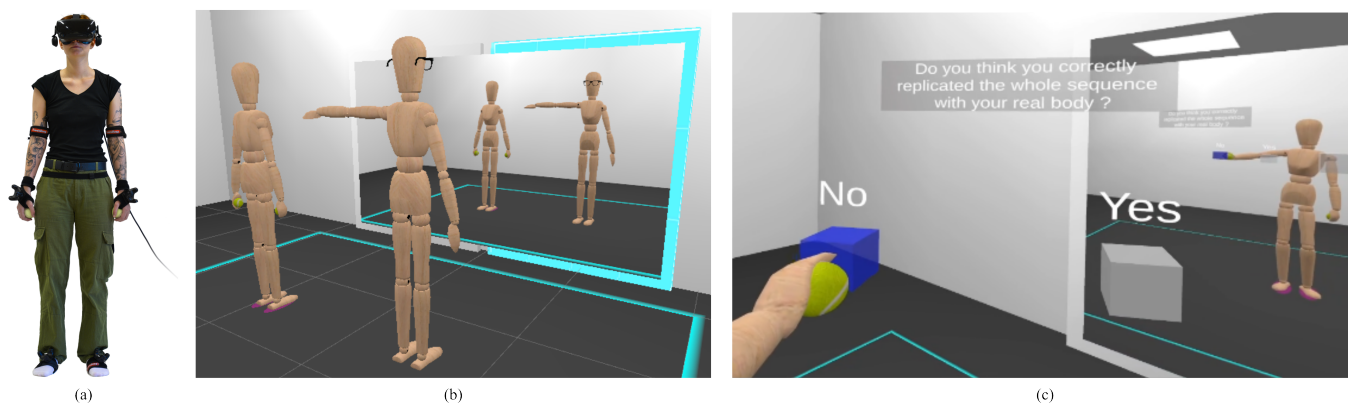


Fig. 1: Experimental set-up. (a) Participant wearing a HMD (Head Mounted Display) and HTC Vive Trackers on both hands, elbows, feet and at hips level; (b) Experimental scene: the participant's avatar (left) and the virtual experimenter showing the movement to perform (right) in front on the full-length mirror; (c) Performance Judgment questionnaire answered in VR; presented after each trial of the main experimental block, participants have to answer "yes" or "no" using the provided virtual buttons.

participants using pre-recorded movements and voice samples, selected in real-time by the real experimenter following a Wizard-of-Oz set-up. The virtual scene, a large empty room with white walls and a dark floor, featured a full-length virtual mirror facing the participant (Figure 1). The mirror frame was split in two, the left part corresponding to the participant's mirror side and the right part to the virtual experimenter's. Both could individually be switched on and off, glowing in a blue color to indicate the state of the task.

### 3.2 Avatar design and full-body motion capture

A gender-neutral and non-realistic wooden mannequin was selected for both the participant's avatar and the virtual experimenter (Figure 1b). This choice offers several advantages, such as avoiding a potential uncanny valley effect [37] while effectively competing with more realistic avatars when considering embodiment [21, 34], as well as minimizing the noticeability of potential motion artifacts [1].

The participant's avatar was animated using the VRIK solver from the FINALIK Unity package (<https://assetstore.unity.com/packages/tools/animation/final-ik-14290>). Such Inverse Kinematic (IK) algorithm computes the avatar's body pose from the participant's physical body tracking data. Here the participant's real body motion was acquired through the use of seven HTC Vive trackers placed on the participant's hands, elbows, feet, and hips (Figure 1a). To account for inaccuracies in the avatar's skeleton structure and to ensure a correct avatar pose, two types of corrections were implemented on the raw tracking data. First, the height of the head and hips IK targets were clamped to the corresponding heights measured during the initial calibration T-pose performed by the participant. Second, if necessary, the length of the legs and arms were automatically adjusted in real-time to prevent the avatar's feet from interpenetrating with the virtual floor and ensure that the virtual arms of the avatar would be fully extended when the participants stood with their arms alongside their body.

### 3.3 Task and experimental conditions

The experimental task consisted in i) memorising a sequence of movements demonstrated by the virtual experimenter and ii) repeating immediately after the whole sequence in front of the mirror. The sequences of movement were inspired from previous work about serial recall of movements [46, 48] and were made of consecutive simple gestures following each other in a pseudo-random order (Figure 2) :

- left or right knee flexed at 90°
- left or right arm raised to shoulder height to the side
- left or right arm raised above head
- both arms raised above head
- left or right arm raised to shoulder height to front of body

- both arms raised to shoulder level to front of body

Each individual movement beginning and ending in the *neutral pose* (standing with the arms relaxed alongside the body, Figure 2.A), there is no constraint for linking them together and they can be freely ordered to generate a large number of different sequences.

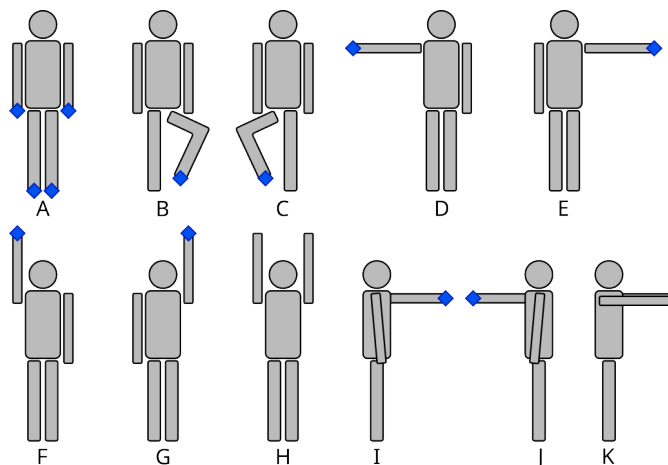


Fig. 2: Movements composing a sequence to remember start from the *neutral posture* A, followed by reaching one of the target postures (B to K), and a return to A. The sequences of movement are thus defined as a succession of postures starting with A, followed by e.g. B, A, H, A, C, etc., and always ending with A. The subset of *reversible* movements (for which the order is reversed in a pair) is: B and C: left / right knee flexed at 90°; D and E: left / right arm raised to shoulder height to the side, F and G: left / right arm raised above head, I and J: left / right arm raised to shoulder height to front of body, G and I: left arm raised above head / to shoulder height to front of body, F and J: right arm raised above head / to shoulder height to front of body, E and I: left arm to shoulder height to the side / to front of body, D and J: right arm to shoulder height to the side / to front of body. Blue markers indicate the reference points used for the automatic classification of participant's movements.

Each trial thus followed the following steps. First, the virtual experimenter demonstrated a sequence of movements in front of the mirror. This was indicated by the blue frame around the right side of the mirror (Figure 1b). The task is made more difficult by asking participants to perform a concurrent motor task during the demonstration of the sequence to remember [46]. This concurrent task consisted in touching both shoulders with the hands, joining hands together at belly level, and touching the hips, and so on until the end of the sequence demon-



stration. During this phase, the left side of the mirror was hidden so participants could not see their reflection. After this demonstration, the left side of the mirror was uncovered and the corresponding frame was highlighted in blue, indicating to the participant that the next phase started. A fixation cross appeared on the mirror at the participant's side (at the level of knee reflection height, so that participants would better see their own hands and body in first person view when performing the movements), and participants were instructed to look at it. After 1 second, the cross disappeared and a simple audio cue was played to instruct participants to start replicating the movements they were shown. When performing the sequence of movements, participants were looking at the reflection of their avatar's body in front of them. The timing of the execution of each movement was imposed; participants had to wait until an audio cue was played before performing the next movement of the sequence (technically, the system detected if participant returned to the neutral posture and waited 0.75s before playing the sound). In addition, to prevent participants from performing the sequence at a faster pace during the recall (natural tendency to finish quickly before forgetting), they were asked to match as much as possible the speed of the original movement.

The experimental manipulation then consisted in altering the movement of the avatar in half of the trials. In the *congruent* condition, no alteration was made and the avatar body followed the participant's tracked body (subsection 3.2). In the *incongruent* condition, the avatar starts by doing the same as the participant but the last two movements of the sequence were replaced by pre-recorded movements (i.e. independent from the actual participant's motion). Of note, the participant's head was always kept tracked during the experiment to avoid motion sickness. The reason for altering the avatar's movement only at the end of the sequences is to take advantage of the difficulty to remember the last items in a long sequence of movements. Previous work indeed indicates a primacy effect but no recency equivalent for the serial recall of movements (i.e. decreasing likelihood of correct recall of an item with its position in the sequence) [48]. With this increased likelihood that participants would not remember what movement to perform at the end of the sequence, we aim at increasing the possibility that participants would be incited to do what the avatar is doing. For this reason, the movements performed by the avatar near the end of the sequence should be different from the correct movements to perform, but similar enough to be mistaken for the correct ones. As previous works on recall of movements [46] have shown that the recall performance is influenced by the similarity in body part, direction (e.g. forward vs. side), and nature of the movement (e.g. flexed or extended limb movement), we considered that the movements to be substituted by the avatar should be similar in at least two aspects. We thus identified couples of *reversible* movements, differing either from the body side (e.g. right vs left arm or leg), direction (e.g. front or side), or amplitude (e.g. shoulder level or above head). For instance, *raising the right arm above the head* and *raising the left arm above the head* are considered *reversible* (only the side changes), while *raising the right arm above the head* is not *reversible* with *raising the left arm to shoulder height to the side* (both side and amplitude would be different). A total of 8 pairs of *reversible* were kept for the experiment (see Figure 2).

To sum up, the sequences used in our experiment are generated pseudo-randomly so as to ensure that a movement is never repeated in a sequence and, most importantly, that the last two movements are *reversible*. Consequently, in the *incongruent* condition, the avatar performs pre-recorded animations corresponding to the reversal of what should be done in the correct order. Said otherwise, when a participant reproduces the sequence they were shown and reached the last 2 movements of the sequence, their avatar performed them in reversed order. Technically, the avatar replayed animations were recorded during a preliminary training phase (see Figure 3).

### 3.4 Experimental design and procedure

Our main experimental block consisted in 16 trials, 8 *congruent* and 8 *incongruent*, presented in a pseudo-random order. To ensure a uniform difficulty of the task for all participants, the experiment started with a sequence length calibration block aiming to determine each

participant's memory performance (Figure 3). The main experimental block then followed using the established sequence length, and added a subjective evaluation of the sense of embodiment after each trial.

**Sequence length calibration phase (Block 1).** In order to find the maximum number of movements each participant can recall correctly, they underwent a first block consisting in repeating sequences of movements of increasing length, starting with  $L = 3$  and ending when the participant systematically failed. The trials in this block followed the same general structure as the main experimental block, but with some noticeable differences. First, the avatar's animation was always *congruent* with the participant's physical movements. Second, participants were not asked to answer a questionnaire after each trial. The criteria for continuing to an increased sequence length were as follows. If the sequence was correctly recalled and if its duration matched the demonstrated one's (with a 20% tolerance), the trial was considered as a *success*. Else, if the participant's recall was either too short or too long, the trial was considered as invalid and another one was performed with a new sequence. Finally, if there was one or more errors in the recalled sequence, the trial was considered as *failed*. For a given length  $L$ , if two out of three trials were considered as *success*, the tested length of sequence was incremented. Otherwise, a verification phase started, which consisted in three additional trials of the same length  $L$ . This verification phase allowed to ensure the considered sequence length to be the participant's actual maximum span, by making sure that a few more training trials would not increase it. Again, if two out of three trials were considered as *success*, the length of sequence tested was incremented and the first block continued. Else, the corresponding length  $L$  was considered as too long for the participant to recall correctly and their individual memory performance limit, denoted  $L_{max}$ , was set at  $L - 1$ .

**Main experimental phase (Block 2).** Based on the sequence length  $L_{max}$  previously established, the main experimental block purposefully challenged participants with 16 different difficult sequences of length  $L_{max} + 2$ . Thus, for the *incongruent* trials, once the participant completed the  $L_{max}^{th}$  movement and returned to the *neutral pose*, their avatar would perform the predefined movement corresponding to the reversed order of the two last movements of the shown sequence (see subsection 3.3). In these cases, the transition from real-time motion capture to the replay of a prerecorded animations was carefully timed to ensure that participants were in the neutral posture at the moment the animation started with the neutral posture. Technically, the movements samples were recorded during the training phase, with each recording starting when the audio cue indicating participants to repeat the movement they were shown was played. Thus, for incongruent trials, the same audio cue was used as a time cue for triggering the replay. More precisely, 0.25s after the validation of the participant's neutral pose, the transition with the replayed movement was ensured with a 0.5s linear interpolation phase to smooth the postural differences. This process duration corresponds to the 0.75s delay before the audio cue was played to indicate the participant to perform the next movement of the sequence, and ensures the avatar's movement to start at the correct moment of the participant's recall. As in the previous block, a trial was considered valid if the duration of the participant's sequence matched the duration of the sequence to repeat, with a 20% tolerance. Else, the trial was invalid and a similar trial was rescheduled later in the block. At the end of every valid trial, participants were asked two subjective ratings and one yes/no question, presented in random order. These questions were presented within the VR environment and the participants used virtual sliders and buttons to provide their answers (Figure 1c). To evaluate participants' subjective experience of body ownership for the virtual body and of agency for the avatar's movement, participants were asked to rate the following affirmations (taken from [22, 28], and used in our previous work [9]) using continuous sliders ranging from "strongly disagree" to "strongly agree"; "*During the whole sequence, I felt as if the virtual body was my body*" and "*During the whole sequence, the virtual body moved just like I wanted to, as if it was obeying my will*". To assert the participants'

judgment of their performance, they were asked to answer *yes* or *no* to the question "Do you think you correctly replicated the whole sequence with your real body?".

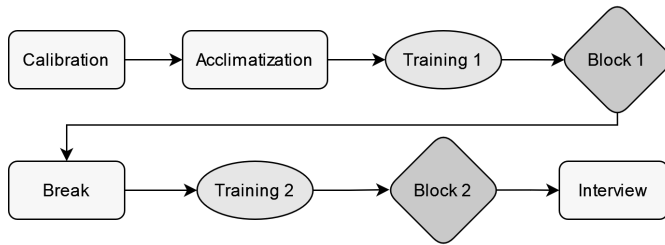


Fig. 3: Experiment structure. The experiment starts with the avatar calibration ( $\approx 2$ min), followed by a small acclimatization phase ( $\approx 30$ s) and the first training phase ( $\approx 10$ min) during which the participants' movements were recorded to be replayed later in the block 2. The block 1 followed, allowing to determine the participant's individual memory span  $L_{max}$  ( $\approx 25$ min). Participant could then take a break ( $\approx 5$ min) before going through the second training phase to familiarize with the procedure of the second block and to the questionnaire ( $\approx 5$ min). The block 2 followed (main experimental block), in which participants were asked to successively recall 16 sequences of length  $L_{max} + 2$  ( $\approx 45$ min). Finally, a small post-experiment interview was conducted ( $\approx 5$ min).

**Overall procedure.** The experiment unfolds according to the following procedure (Figure 3). Upon their arrival, participants performed a quick balance test (inclusion criteria), signed a consent form and filled a demographic questionnaire. They were equipped with seven Vive trackers, a Valve Index HMD, and given two tennis balls to hold in their hands (this prevented finger movements and visually matched the hands of the wooden mannequin avatar (Figure 1). After explaining the concurrent motor task and ensuring that the participant could perform it correctly, the VR experiment started. In VR, the virtual experimenter explained how the experiment would unfold and guided the calibration of the avatar: the avatar was globally scaled to match the participant's height, and its arms and legs length were adjusted to fit the participant. Followed a small acclimatization phase during which participants were instructed to look at the virtual experimenter doing simple movements in front of the mirror and to repeat them. In the training phase that followed, participants familiarized themselves with the mirror set-up, the fixation cross and the audio cue, and trained to perform the 10 movements composing the sequences as well as a short sequence of two movements (see subsection 3.4). Of note, these movements were recorded so as to be replayed later in the main experimental block. Participants then performed the sequence length calibration block (block 1). When preliminary piloting the task, we had established that the experiment would be too long for participants in some rare cases when participants performed exceptionally well. We thus decided to exclude outliers demonstrating exceptional memory performances of sequences of seven or more movements (previous work showing that the average performance is of 3.43 movements to recall,  $std \approx 0.7$  [46]). Participants could then take a break before continuing with a quick training phase conducted to get them used to the procedure of the main experimental block (example recall of two sequences with  $L = 2$ ) and to the questionnaire (see subsection 3.4). Importantly, the movements of the avatar were congruent with the participant's ones during the first sequence recall, but were altered during the second recall (the avatar's movements were frozen), to show participants that their avatar could sometimes have a different behavior and to give them a reference point for the questionnaire rating. The main experimental block (block 2) followed, in which both the tracking data of the participants and the movements of their avatar were recorded for further analysis. Eventually, the virtual experimenter thanked the subjects for their participation, and told them they could remove the equipment. Finally, once the HMD and Vive trackers were removed, a small post-experiment interview was conducted by the real experimenter (see supplemental materials).

### 3.5 Hypotheses and data pre-processing

As a prerequisite to our manipulation, it was expected that participants would report a high level of embodiment for their avatar when their movements are *congruent* (replicating [9]). Then, because we experimentally introduced a rather large discrepancy between the participants' and the avatar's movements, we expected the movement congruency factor to be associated with a significant change in the reported embodiment, with both a lower sense of body ownership and agency for the avatar's movement in *incongruent* trials than in *congruent* ones (H0). Second, we hypothesized that participants would be prone to do the same movement as their avatar at the end of the recall of long sequences of movements, although their avatar would perform the wrong movements (i.e. when the avatar is reversing the order of the two last movements in *incongruent* trials). To verify this, we evaluated whether participants performed partial or total movement reversals more often in *incongruent* trials than in *congruent* ones (H1). In addition, and because there is a lower likelihood that the end of a sequence is recalled correctly after an error has been made [48], we expected a larger tendency to perform the same movement as the avatar when errors already occurred in the recall before the last two movements of the sequence (H2). Together, these two hypothesis (H1 and H2) would confirm that participants have a spontaneous tendency to follow what their avatar is doing when they are unsure about what movement to perform. Said otherwise, this would indicate that the avatar is somehow guiding the participants to perform movements in case of uncertainty. Finally, to explore the link between the subjective experience of embodiment and the tendency to follow their self-avatar, it can be hypothesized that participants would rate higher their sense of agency for the avatar's movement and of ownership in *incongruent* trials in which they did the same movements as their avatar as opposed to when they didn't (H3). This could indeed possibly mean that participants experienced a weaker disruption between their real and virtual bodies after having corrected for a multi-sensory discrepancy. The alternative would rather be in favor of an acceptance of a clearly perceived discrepancy, breaking the sense of ownership of the virtual body and agency towards the avatar's movements.

**Movement data pre-processing.** The validation of H1 and H2 required analyzing participants recalls for both *congruent* and *incongruent* trials of the main experimental block. This analysis required to detect if participants did the same movements as their avatar in *incongruent* trials, and to distinguish such behavior from recall errors. For instance, for a sequence ending with the movements M1 then M2 (Figure 4), the participant may end the recall with movements M2 and M1, which would simply indicate an error in the recall if the trial is *congruent*. The probability of such spontaneous *total user swap* of the two last movements when recalling a sequence can be measured by analysing the *congruent* trials. Conversely, in the *incongruent* trials, the avatar reversed the order of the last two movements, replaying M2 first and M1 afterwards. In this situation, a participants also ending the sequence with movements M2 and M1 can either correspond to a simple recall error, but can also indicate a *self-avatar follower effect*. Therefore, the occurrence of the *self-avatar follower effect* would be confirmed by a significant increase in the number of occurrence of *user swaps* (corresponding to what the avatar is showing) in the *incongruent* trials as compared to the *congruent* trials. All movements composing the recalls should therefore be classified in all trials to identify trials in which a *user swap* occurred for the last two movements (either a total inversion or a partial swap resulting from a mixing with another movement). This analysis thus first requires to compare what had to be replicated with what was actually performed by the participant, for each movement of all sequences (Figure 4).

**Participant's movements classification.** To classify each participant's movement performed during the recalls, we first determined a set of *reference points* characterizing the movements composing the sequences (Figure 2). To do so, we considered the single-limb movements samples recorded during the first training phase as their corresponding label is known by design. Because every movement

composing the sequences consists in a raising motion of either a hand (or both hands) or a foot, we focused on the  $y$  coordinate (vertical axis) of the movements data. Indeed, for a given movement, the highest point reached by the participant's corresponding hand or foot coincides with the apex of the movement, achieved before going back to the neutral pose. The associated 3D position  $(x_{ref}, y_{ref}, z_{ref})$  of the moving limb apex therefore defined the *reference point* of the considered movement. We then defined 12 *reference points* (see blue markers on Figure 2) corresponding to both resting and raised positions of hands and feet, allowing to characterize all the movements involved in the sequences. Once this set of *reference points* was computed for each participant, each movement performed during the recall of the sequences in the second block could be classified according to the following methodology. For a given movement of a sequence recall, hands and feet motions were first classified independently before being combined to infer the corresponding full-body pose. To do so, similarly to the computation of the *reference points*, the highest positions reached by each hand (resp. foot) were identified and compared to the associated four (resp. two) *reference points*. The corresponding closest *reference points* were then determined using euclidean distance, thereby associating to each hand and foot a *reference point*. Finally, the resulting four classifications were combined to deduce the full-body motion. For instance, for a given motion sample, if both feet were associated to their *neutral pose reference point*, and both hands to their *raised above head reference point*, the corresponding motion sample was labeled as *both arms raised above head*.

**Trials labelling.** Once each movement of the recalls was classified, each trial of the second experimental block was labeled according to the pattern formed by the last two movements of the reference sequence and of the associated recall (Figure 4). First, if the reference sequence ended with the movements M1 then M2, if the last two movements performed by the participant during the recall were M2 then M1, the trial was labelled as *total user swap*. Else if the participant ended the sequence with a movement X that was not M2 followed by the movement M1, or conversely by the movement M2 followed by a movement Y that was not M1, the trial was labeled as *partial user swap*. Finally if the participant ended the sequence with either a correct recall, or a couple of movements not corresponding to the previous cases, the trial was labeled as *no user swap*.

## 4 RESULTS

36 healthy volunteers were recruited for this study. Four were excluded because of technical issues during the experiment, and five after the first block as they could recall correctly sequences of seven movements or more. The statistical analysis was thus conducted on 27 subjects (17 females), aged from 18 to 36 (mean = 22.6, std  $\approx$  3.4). All participants gave informed consent and the study was approved by our local ethics committee. The study and methods were carried out in accordance with the guidelines of the declaration of Helsinki. The experiment lasted between one hour and a half and two hours, depending on the length of the movement sequences subjects could correctly remember.

Over the 432 trials composing the dataset, 17 trials were excluded during the pre-processing of the data (resp. 3 *congruent* and 14 *incongruent* trials). Among them, four trials were excluded because of technical issues preventing their correct completion, eight due to participants performing two movements instead of one between two audio cues, one as a result of the participant not moving during the avatar's replay, and four because a participant's recall movement could not be classified (typically not corresponding to any of the ten reference movements). Therefore, 415 trials were included in the analysis, the resulting dataset containing 213 *congruent* trials and 202 *incongruent* trials.

### 4.1 Effect of movement congruency on embodiment

Concerning the effect of the avatar movements congruency on the ratings of sense of agency for the avatar's movement and of ownership for the avatar's body, we compared the corresponding scores between *congruent* and *incongruent* trials. As the samples are unbalanced,

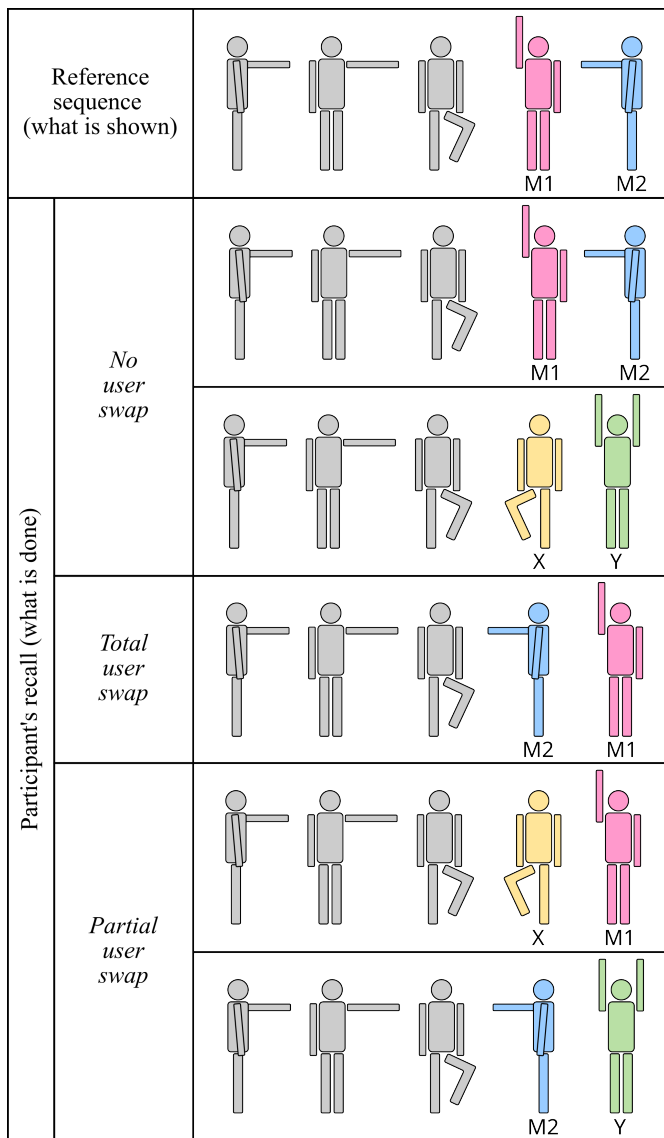


Fig. 4: Trials labelling. The first row shows an example sequence of five movements to reproduce, ending with movements M1 and M2. A participants of our experiment might have recalled perfectly, failed entirely or, importantly, might have combined M1 and M2 in the opposite order and/or mixed with other movements. Because in our manipulation in *incongruent* trials the avatar reversed the order of the two last movements, trials were labeled as *no*, *total* or *partial user swap* in order to evaluate if the avatar could have influenced participant when performing their recall.

two-sided permutation tests were applied. Both tests were significant ( $p < 0.001$ ), showing the sense of agency for the avatar's movement and ownership for the avatar to be significantly higher when its movements were *congruent* than *incongruent* with the participants' ones (Figure 5), thereby validating our first hypothesis (H0). Of note, the average of both scores in *congruent* trials is  $\approx 0.84$ , thus confirming a high embodiment for the avatar in non-manipulated conditions.

### 4.2 Task performance and occurrence of the follower effect

On average, based on the individual memory span determined during the first experimental block, participants could recall correctly sequences of length  $L_{max} \approx 4.37$  movements (std  $\approx 0.95$ ), which is in line with previous observations [46]. Participants' recall performance in the main experimental block was determined based on the movements classification performed for each trial. Participants failed to recall the sequences correctly in approximately 80% of the



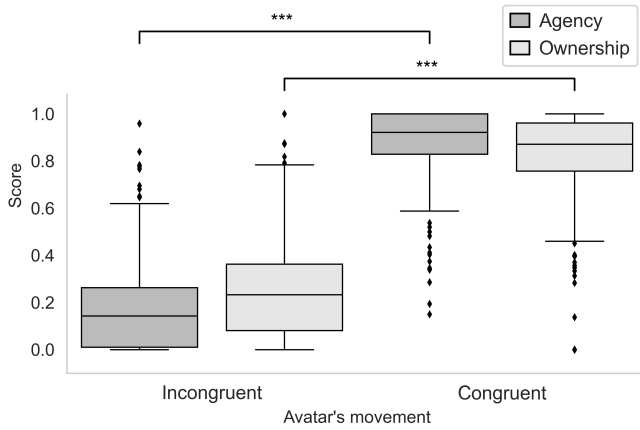


Fig. 5: Subjective experience of body ownership for the virtual body and of agency for the avatar's movement depending on avatar's movement congruency. Both scores range from 0 ("strongly disagree") to 1 ("strongly agree"). For *incongruent* trials,  $m_{agency} \approx 0.18$ ,  $std_{agency} \approx 0.19$ ,  $m_{ownership} \approx 0.27$ ,  $std_{ownership} \approx 0.23$ . For *congruent* trials,  $m_{agency} \approx 0.88$ ,  $std_{agency} \approx 0.16$ ,  $m_{ownership} \approx 0.81$ ,  $std_{ownership} \approx 0.23$ .

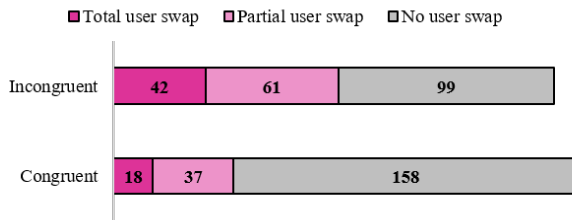


Fig. 6: User swaps repartition across the 213 *congruent* and the 202 *incongruent* trials. *Total user swaps* represent  $\approx 21\%$  of the *incongruent* trials and  $\approx 8\%$  of the *congruent* trials. *Partial user swaps* represent  $\approx 30\%$  of the *incongruent* trials and  $\approx 17\%$  of the *congruent* trials. Taken together, *user swaps* represent  $\approx 51\%$  of the *incongruent* trials and  $\approx 26\%$  of the *congruent* trials.

trials (85 correct recalls out of 415 trials in total), confirming that our procedure produced hard-to-recall sequences. Comparing participants' success judgment to their actual performance, participants mainly answered *no* to the performance judgment question when they actually failed to recall the sequence correctly (257 *no* out of 330 trials showing an incorrect recall). Conversely, in case of a correct recall (85 trials), participants mainly answer *yes* (77 trials). Therefore, with around 80% of judgment accuracy, participants were quite accurate when judging their recall performance.

To estimate the occurrence of the *self-avatar follower effect*, we first compared the different types of *user swap* frequencies depending on the avatar motion condition (Figure 6). Considering both *partial* and *total user swaps* together, 103 out of 202 *incongruent* trials, and 55 out of 213 *congruent* trials, exhibited a *user swap*. The associated Pearson's Chi-squared test was significant ( $p < 0.001$ ), showing the avatar's movement congruency to have an effect of the *user swap* occurrence. Furthermore, 42 over 202 *incongruent* trials, and 18 over 213 *congruent* trials showed a *total user swap*, with the corresponding Pearson's Chi-squared test being also significant ( $p < 0.001$ ). Taken together, results show a significant difference in the occurrence of *user swaps* between *congruent* and *incongruent* trials, thereby validating our hypothesis (H1).

Second, to determine whether there is a larger tendency to follow the avatar when errors were already made in the recall before the avatar's movements manipulation, we compared the occurrence of *user*

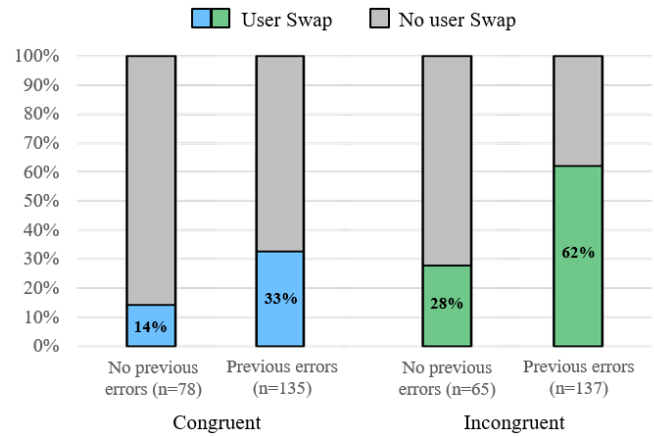


Fig. 7: User swaps repartition depending on the occurrence of errors before the last two movements of the sequences. Among the 213 *congruent* trials,  $\approx 14\%$  of the trials not exhibiting errors in the recall before the last two movements of the sequences ended with a *user swap*. This proportion doubles ( $\approx 33\%$ ) among trials exhibiting errors in the recall before the last two movements of the sequences. Similarly, among the 202 *incongruent* trials,  $\approx 28\%$  of the trials not exhibiting errors in the recall before the last two movements of the sequences ended with a *user swap*. This proportion doubles ( $\approx 62\%$ ) among trials exhibiting errors in the recall before the last two movements of the sequences. Additionally, focusing on trials with errors before the last two movements of the sequence, we observe around twice more *user swaps* in *incongruent* trials than in *congruent* trials. Similarly, in trials without errors before the last two movements of the sequence, there are twice more *user swaps* in *incongruent* trials than in *congruent* trials.

*swaps* depending on the correctness of the recall before the last two movements of the sequences (Figure 7). Over the 213 *congruent* trials, 135 trials exhibited errors in the recall before the last two movements of the sequences. Among them, 44 showed a *partial* or *total user swap*. Additionally, over the 78 trials without errors in the recall before the last two movements of the sequences, only 11 trials exhibited a *partial* or *total user swap*. Similarly, over the 202 *incongruent* trials, 137 trials exhibited errors in the recall before the last two movements of the sequences. Among them, 85 showed a *partial* or *total user swap*. Additionally, over the 65 trials without errors in the recall before the last two movements of the sequences, only 18 trials exhibited a *partial* or *total user swap*. For both *congruent* and *incongruent* trials, the associated Pearson's Chi-squared test were significant (resp.  $p \approx 0.003$  and  $p < 0.001$ ), showing the distribution of *user swaps* to depend on the occurrence of previous errors in the recall. Indeed, in both *congruent* and *incongruent* trials, we observe twice more trials exhibiting a *partial* or *total user swap* in trials in which a recall error happened before the last two movements of the sequence than when no such error occurred. Our hypothesis (H2) is thus only partially validated as the presence of previous errors in the recall indeed influences the occurrence of *user swaps* at the end of the sequence, but in the same proportions for *congruent* and *incongruent* trials. However, for both trials with and without errors before the last two movements of the sequence, we observe around twice more *user swap* in *incongruent* trials than in *congruent* trials (associated Pearson's Chi-squared tests are significant with resp.  $p < 0.001$  and  $p \approx 0.04$ ), further highlighting the effect of the avatar replay on the occurrence of such *user swap*.

### 4.3 Link between follower-effect occurrence and embodiment scores

To determine if the embodiment scores would be related to the occurrence of a coordination of movement between participants and their avatar, we first compared the scores in *incongruent* trials exhibiting a *total user swap* with those showing *no user swap*. As the samples to compare are unbalanced (resp. 42 and 99 trials), two-sided permutation

tests were applied. Neither the sense of ownership nor the sense of agency toward the avatar's movements exhibited a significant difference (resp.  $p \simeq 0.61$  and  $p \simeq 0.14$ ). Second, when comparing the *incongruent* trials exhibiting only a *partial user swap* to the *incongruent* trials showing *no user swap*, the corresponding two-sided permutation tests were both non significant (resp.  $p \simeq 0.36$  and  $p \simeq 0.63$ ), thus not validating the (H3) hypothesis.

## 5 DISCUSSION

Our experimental setup with immersive VR is overall validated, with very high scores of embodiment in absence of experimental manipulation of the avatar movement (Figure 5). This is valid both for the sense of agency for the avatar's movement, and for the sense of ownership for the avatar's body. Also, as hypothesized, disconnecting the avatar movement from the participants' control is breaking embodiment [31, 39, 40, 42], as shown by a significant difference in embodiment score between *incongruent* and *congruent* trials.

In this context, the experimental task participants had to perform was specifically aiming at provoking moments of doubts and uncertainty about which movement to perform. We first confirmed that participants were at the limit of their recall ability, as our data confirm the occurrence of recall error in  $\simeq 80\%$  of trials. Specifically, we observe, as normally expected, spontaneous errors at the end of the recalls when no movement manipulation is applied, with  $\simeq 63\%$  of the *congruent* trials showing errors in the recall of the last two movements of the sequences. Such recall errors can be about forgetting which movements to perform, or mixing the order of the items. In some cases, participants could have swapped the order of the last two movements, or mixed one of them with another movement. Then, at these crucial moments, we performed our experimental manipulation wherein the avatar would not follow the participants' movement anymore, but would perform a movement on its own, i.e. swap the order of the last two movements. With twice more *user swaps* in *incongruent* trials than in *congruent* trials ( $\simeq 51\%$  against  $\simeq 26\%$ ), we confirm that our manipulation had a strong influence on the participants' recalls. Additionally, our results point toward a larger propensity to perform a *user swap* in trials showing errors in the recall before the last two movements of the sequences than in trials showing no previous errors. Indeed, making a previous recall error increases the odds of making another at the end of the sequence, and in particular the odds of making a *user swap*. However, such effect is of the same magnitude for *congruent* and *incongruent* trials, with around twice more *user swap* when errors were already made before the last two movements of the sequence. Importantly, considering either trials with or without errors before the last two movements of the sequence, we also observe twice more *user swap* in *incongruent* trials than in *congruent* trials, further highlighting the effect of the avatar replay on the occurrence of such *user swap*. Taken together, these two behavioral observations allow us to conclude that our experimental manipulation successfully took advantage of the difficulty of the task to induce a follower-effect toward the self-avatar (Figure 6, Figure 7).

Moreover, and importantly for interpreting these results, we observe that, in *incongruent* trials, participants did not rate differently their subjective experience of embodiment in situations where they followed their avatar than when they didn't. More specifically, they rated low their experience of agency for the avatar's movements (similarly to the "instantaneous" experimental condition in [25]), thus indicating that they noticed that the avatar moved by itself. Still, despite experiencing the dissociation from their avatar, participants nonetheless tended to perform the movement showed by their avatar.

The post-experiment interviews give us some interesting leads about the motivations behind this behavior. To the question *Did you feel like your movements were sometimes influenced by something out of your will?*, 19 out of 27 participants answered positively (i.e.  $\simeq 70\%$ ) and when asked *Did you ever follow the avatar's movement although you wanted to perform another one?*, 20 participants indicated they did at some point (i.e.  $\simeq 74\%$ ). When requested to elaborate, several participants mentioned an impulse to do the same as the avatar when it was moving by itself, e.g., *"When I was seeing the avatar doing a*

*movement, I wanted to do the same"*. Further elaborations about why they would follow the avatar give more precision. At first, linked to the difficulty of the task, it appears that the situation of uncertainty and the possibility that the avatar might be right probably played a role, such as expressed in comments like *"I was doubting so I followed the instinct to follow"*, or *"I was not sure so I followed him because I thought it was the right answer"*. Worth noticing, one participant indicated having followed the avatar despite being sure it was not the correct movements to perform; *"I followed even when I knew it was not the right things, very unconsciously"*. Secondly, linked to the expectation of continuity of embodiment, participants' reports indicate a will to be one with the avatar, to maintain the connection; *"I just wanted to be in a connection with him"*, *"I thought the avatar was me so I did the same movement as the avatar"* or, *"I followed because he was following me, "because my body" you see, so I followed him too"*. Put together, the feedback from participants complement our experimental observations and tend to indicate that, despite their awareness about both the avatar movement disruption and on the possible influence it had on their choice, following the avatar seemed to be the natural thing to do.

**Limitations** Proving the existence of conditions favoring specific behaviors does however not explain the cognitive mechanism behind them. We successfully demonstrate that the *self-avatar follower effect* is not limited to movement distortions that still partially reflects the user's movement; however several limitations in our experimental design prevent us from fully understanding it.

First, and by design, our participants evaluated their subjective embodiment at the end of the trials, right after the avatar movement disruption: the dissociation could not be ignored and was thus reported. It cannot however be concluded that a break of embodiment would have spontaneously been reported in a more ecological context where, for instance, the task would be continued with congruent avatar animation. Nor can it be concluded that embodiment couldn't be restored very rapidly if the disruption itself was of short duration; the avatar could attempt at showing a possible movement option, and if followed restore the control to the user. Similarly, because we ask participants after each trial if they were correct for the whole sequence, we cannot know if they specifically noticed the inversion of the last two movements. This prevent us from generalizing that they followed the avatar despite being aware of the swap in particular, and more specific questions could be asked instead.

Second, our design is also limited in the way it creates a situation of uncertainty about which movement to perform. To ground our experiment on known behavioral effects, we relied on the probability to invert the order of the last two elements of a long sequence to remember. The constraints this imposed to the task and to the analysis complicated its interpretation, in particular regarding the different subjective importance that can have a *correct* or an *incorrect* answer at recall. There could however be many alternative situations to investigate, with different uncertainty factors such as when presented with two-alternative forced choice or a free choice. Both opposite hypotheses could thereby be tested, whether a self-initiated decision would be associated with a strong agency and a rejection of the avatar's guidance, or alternatively, if in the absence of specific motivation for one or the other option, a suggestion by the avatar would be welcome.

Third, our demonstration is based on the observation that participants were influenced by their avatar in a way that led them to do more often a specific type of mistake. The opposite manipulation, where the last two movements would be replayed in the correct order, would be helping the participants and should therefore lead to better performance for the last two movements recall in trials with avatar's movement manipulation than without manipulation. Based on previous work comparing avatar manipulations for helping or hindering movement [17, 24, 40], the *self-avatar follower effect* might even be more likely to occur in such case. Such system could find an application for assisting VR users in performing complex tasks, such as learning series of gestures (e.g., surgical procedures, aerospace training), or to guide them through different exercises (e.g., reeducation sessions).

Finally, the present experiment does not investigate if and how the



feedback of the self-avatar is different from the influence of viewing an agent or another avatar. Indeed, it is possible that other forms of visual guides, such as seeing other participant's avatars or agents performing the task, would show a similar influence on participant's movements. Such comparison would allow to quantify the relative importance of such external influence, and help us understand the specificity of the influence of the self-avatars.

## 6 CONCLUSION

When immersed in VR, users can experience a strong link with their virtual counterpart, their avatar. This feeling of embodiment depends on how the avatar follows the will and mimics the user's physical body. However, it was also observed that, in some cases of small or progressive temporal and spatial manipulations of the avatars' movements, users may spontaneously and reciprocally follow the movement shown by their avatar. Our experiment demonstrates that a similar *self-avatar follower effect* can also be observed when the avatar disconnects and replays a pre-recorded animation instead of following the user. More specifically, our results show that participants tended to do the same movement as their avatar when they were unsure of which movement they should perform. Despite the several improvements that could be made to our paradigm, it demonstrates the possibility of systematically observing such behavior and therefore, contributes to the understanding of the complex interactions between users and their self-avatar. The self-avatar follower effect can therefore be understood as a general behaviour where users are adapting their movement in a way that leads to a reduction in the discrepancy between self-body and seen avatar body, might they be aware or not of their behaviour. Further investigation should be conducted to better characterize if and how the feedback of the self-avatar is different from the influence that can have the view of another character or another person, but it is certain that the possibility of mixing and alternating between following and guiding the user is a feat that only a virtual self-avatar can do, and that this feature should be investigated for improved human-computer interaction in VR.

## SUPPLEMENTAL MATERIALS

All supplemental materials are provided with the manuscript submission. In particular, they include (1) csv files containing the data for the analyses, (2) analysis scripts, (3) a video showing the different experimental conditions, (4) the post experiment interview.

## ACKNOWLEDGMENTS

This work was supported by the Fonds National Suisse de la Recherche Scientifique under the grant 200020\_207424.

## REFERENCES

- [1] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE Virtual Reality (VR)*, pp. 3–10, Mar. 2016. ISSN: 2375-5334. doi: [10.1109/VR.2016.7504682](https://doi.org/10.1109/VR.2016.7504682) 3
- [2] T. Asai. Illusory body-ownership entails automatic compensative movement: for the unified representation between body and action. *Exp Brain Res*, p. 9, 2015. 2
- [3] R. C. Atkinson and R. M. Shiffrin. Human memory: A proposed system and its control processes. In *Psychology of learning and motivation*, vol. 2, pp. 89–195. Elsevier, 1968. 2
- [4] M. Azmandian, M. Hancock, H. Benko, E. Ofek, and A. D. Wilson. Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 1968–1979. ACM, San Jose California USA, May 2016. doi: [10.1145/2858036.2858226](https://doi.org/10.1145/2858036.2858226) 1
- [5] A. D. Baddeley and G. Hitch. Working memory. In *Psychology of learning and motivation*, vol. 8, pp. 47–89. Elsevier, 1974. 2
- [6] O. Blanke. Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13(8):556–571, 2012. 1, 2
- [7] O. Blanke and T. Metzinger. Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences*, 13(1):7–13, Jan. 2009. doi: [10.1016/j.tics.2008.10.003](https://doi.org/10.1016/j.tics.2008.10.003) 1
- [8] O. Blanke, M. Slater, and A. Serino. Behavioral, Neural, and Computational Principles of Bodily Self-Consciousness. *Neuron*, 88(1):145–166, Oct. 2015. doi: [10.1016/j.neuron.2015.09.029](https://doi.org/10.1016/j.neuron.2015.09.029) 1, 2
- [9] L. Boban, L. Strauss, H. Decroix, B. Herbelin, and R. Boulic. Unintentional synchronization with self-avatar for upper-and lower-body movements. *Frontiers in Virtual Reality*, 4:1073549, 2023. 1, 2, 4, 5
- [10] M. Botvinick and J. Cohen. Rubber hands ‘feel’ touch that eyes see. *Nature*, 391(6669):756–756, Feb. 1998. doi: [10.1038/35784](https://doi.org/10.1038/35784) 2
- [11] P. Bourdin, M. Martini, and M. V. Sanchez-Vives. Altered visual feedback from an embodied avatar unconsciously influences movement amplitude and muscle activity. *Scientific reports*, 9(1):19747, 2019. 1
- [12] D. Burin, K. Kilteni, M. Rabuffetti, M. Slater, and L. Pia. Body ownership increases the interference between observed and executed movements. *PLOS ONE*, 14(1):e0209899, Jan. 2019. doi: [10.1371/journal.pone.0209899](https://doi.org/10.1371/journal.pone.0209899) 2
- [13] A. Cheymol, A. Lécuyer, J.-M. Normand, F. Argelaguet, et al. Beyond my real body: Characterization, impacts, applications and perspectives of “dissimilar” avatars in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 1
- [14] B. A. Cohn, A. Maselli, E. Ofek, and M. Gonzalez-Franco. Snapmove: Movement projection mapping in virtual reality. In *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pp. 74–81. IEEE, 2020. 1, 2
- [15] H. G. Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulic. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PLOS ONE*, 12(12):e0190109, Dec. 2017. Publisher: Public Library of Science. doi: [10.1371/journal.pone.0190109](https://doi.org/10.1371/journal.pone.0190109) 2
- [16] H. G. Debarba, E. Molla, B. Herbelin, and R. Boulic. Characterizing embodied interaction in First and Third Person Perspective viewpoints. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 67–72, Mar. 2015. doi: [10.1109/3DUI.2015.7131728](https://doi.org/10.1109/3DUI.2015.7131728) 2
- [17] M. Delahaye, O. Blanke, R. Boulic, and B. Herbelin. Avatar error in your favor: Embodied avatars can fix users' mistakes without them noticing. *Plos one*, 18(1):e0266212, 2023. 8
- [18] H. H. Ehrsson. The Experimental Induction of Out-of-Body Experiences. *Science*, 317(5841):1048–1048, Aug. 2007. Publisher: American Association for the Advancement of Science Section: Brevia. doi: [10.1126/science.1142175](https://doi.org/10.1126/science.1142175) 2
- [19] J. C. Eubanks, A. G. Moore, P. A. Fishwick, and R. P. McMahan. The Effects of Body Tracking Fidelity on Embodiment of an Inverse-Kinematic Avatar for Male Participants. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 54–63. IEEE, Recife/Porto de Galinhas, Nov. 2020. doi: [10.1109/ISMAR50242.2020.00025](https://doi.org/10.1109/ISMAR50242.2020.00025) 2
- [20] C. Farrer, M. Bouchereau, M. Jeannerod, and N. Franck. Effect of distorted visual feedback on the sense of agency. *Behavioural neurology*, 19(1-2):53–57, 2008. 2
- [21] R. Fribourg, F. Argelaguet, A. Lécuyer, and L. Hoyet. Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2062–2072, May 2020. doi: [10.1109/TVCG.2020.2973077](https://doi.org/10.1109/TVCG.2020.2973077) 3
- [22] R. Fribourg, N. Ogawa, L. Hoyet, F. Argelaguet, T. Narumi, M. Hirose, and A. Lécuyer. Virtual Co-Embodiment: Evaluation of the Sense of Agency while Sharing the Control of a Virtual Body among Two Individuals. *IEEE Transactions on Visualization and Computer Graphics*, 27(10):4023–4038, Oct. 2021. 16. doi: [10.1109/TVCG.2020.2999197](https://doi.org/10.1109/TVCG.2020.2999197) 4
- [23] K. J. Friston, J. Daunizeau, J. Kilner, and S. J. Kiebel. Action and behavior: a free-energy formulation. *Biological cybernetics*, 102(3):227–260, 2010. 2
- [24] H. Galvan Debarba, R. Boulic, R. Salomon, O. Blanke, and B. Herbelin. Self-attribution of distorted reaching movements in immersive virtual reality. *Computers & Graphics*, 76:142–152, Nov. 2018. doi: [10.1016/j.cag.2018.09.001](https://doi.org/10.1016/j.cag.2018.09.001) 2, 8
- [25] M. Gonzalez-Franco, B. Cohn, E. Ofek, D. Burin, and A. Maselli. The self-avatar follower effect in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 18–25. IEEE, 2020. 1, 2, 8
- [26] P. Haggard and V. Chambon. Sense of agency. *Current biology*, 22(10):R390–R392, 2012. 2
- [27] B. Herbelin, R. Salomon, A. Serino, and O. Blanke. Neural mechanisms of bodily self-consciousness and the experience of presence in virtual reality. Technical report, De Gruyter, 2016. 1

- [28] A. Kalckert and H. H. Ehrsson. The spatial distance rule in the moving and classical rubber hand illusions. *Consciousness and cognition*, 30:118–132, 2014. 4
- [29] S. Kasahara, K. Konno, R. Owaki, T. Nishi, A. Takeshita, T. Ito, S. Kasuga, and J. Ushiba. Malleable Embodiment: Changing Sense of Embodiment by Spatial-Temporal Deformation of Virtual Human Body. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 6438–6448. ACM, Denver Colorado USA, May 2017. doi: 10.1145/3025453.3025962 2
- [30] K. Kilteni, R. Groten, and M. Slater. The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, Nov. 2012. 1. doi: 10.1162/PRES\_a\_00124 1, 2
- [31] E. Kokkinara and M. Slater. Measuring the Effects through Time of the Influence of Visuomotor and Visuotactile Synchronous Stimulation on a Virtual Body Ownership Illusion. *Perception*, 43(1):43–58, Jan. 2014. doi: 10.1068/p7545 1, 2, 8
- [32] E. Kokkinara, M. Slater, and J. López-Moliner. The Effects of Visuomotor Calibration to the Perceived Space and Body, through Embodiment in Immersive Virtual Reality. *ACM Transactions on Applied Perception*, 13(1):1–22, Dec. 2015. doi: 10.1145/2818998 2
- [33] P. Lanillos, S. Franklin, A. Maselli, and D. W. Franklin. Active strategies for multisensory conflict suppression in the virtual hand illusion. *Scientific Reports*, 11(1):22844, 2021. 2
- [34] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10. ACM, Gothenburg Sweden, Nov. 2017. doi: 10.1145/3139131.3139156 3
- [35] B. Lenggenhager, T. Tadi, T. Metzinger, and O. Blanke. Video Ergo Sum: Manipulating Bodily Self-Consciousness. *Science*, 317(5841):1096–1099, Aug. 2007. Publisher: American Association for the Advancement of Science Section: Report. doi: 10.1126/science.1143439 2
- [36] A. Maselli, P. Lanillos, and G. Pezzulo. Active inference unifies intentional and conflict-resolution imperatives of motor control. *PLoS computational biology*, 18(6):e1010095, 2022. 2
- [37] M. Mori, K. F. MacDorman, and N. Kageki. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine*, 19(2):98–100, June 2012. Conference Name: IEEE Robotics Automation Magazine. doi: 10.1109/MRA.2012.2192811 3
- [38] E. F. Pavone, G. Tieri, G. Rizza, E. Tidoni, L. Grisoni, and S. M. Aglioti. Embodying others in immersive virtual reality: electro-cortical signatures of monitoring the errors in the actions of an avatar seen from a first-person perspective. *Journal of Neuroscience*, 36(2):268–279, 2016. 1
- [39] T. Porssut, O. Blanke, B. Herbelin, and R. Boulic. Reaching articular limits can negatively impact embodiment in virtual reality. *Plos one*, 17(3):e0255554, 2022. 1, 2, 8
- [40] T. Porssut, B. Herbelin, and R. Boulic. Reconciling being in-control vs. being helped for the execution of complex movements in vr. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 529–537. IEEE, 2019. 1, 2, 8
- [41] T. Porssut, Y. Hou, O. Blanke, B. Herbelin, and R. Boulic. Adapting virtual embodiment through reinforcement learning. *IEEE Transactions on Visualization and Computer Graphics*, 28(9):3193–3205, 2021. 2
- [42] T. Porssut, F. Iwane, R. Chavarriaga, O. Blanke, J. d. R. Millán, R. Boulic, and B. Herbelin. Eeg signature of breaks in embodiment in vr. *Plos one*, 18(5):e0282967, 2023. 1, 2, 8
- [43] M. Rietzler, F. Geiselhart, and E. Rukzio. The matrix has you: realizing slow motion in full-body virtual reality. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10. ACM, Gothenburg Sweden, Nov. 2017. doi: 10.1145/3139131.3139145 1, 2
- [44] N. Sajid, P. J. Ball, T. Parr, and K. J. Friston. Active inference: demystified and compared. *Neural computation*, 33(3):674–712, 2021. 2
- [45] R. Salomon, N. B. Fernandez, M. van Elk, N. Vachicouras, F. Sabatier, A. Tychinskaya, J. Llobera, and O. Blanke. Changing motor perception by sensorimotor conflicts and body ownership. *Scientific Reports*, 6(1):25847, May 2016. Number: 1 Publisher: Nature Publishing Group. doi: 10.1038/srep25847 2
- [46] M. M. Smyth, N. A. Pearson, and L. R. Pendleton. Movement and working memory: Patterns and positions in space. *The Quarterly Journal of Experimental Psychology*, 40(3):497–514, 1988. 2, 3, 4, 5, 6
- [47] M. Tsakiris, G. Prabhu, and P. Haggard. Having a body versus moving your body: How agency structures body-ownership. *Consciousness and Cognition*, 15(2):423–432, June 2006. doi: 10.1016/j.concog.2005.09.004 2
- [48] F. Wachowicz, C. J. Stevens, and T. P. Byron. Effects of balance cues and experience on serial recall of human movement. *Dance Research*, 29(supplement):450–468, 2011. 2, 3, 4, 5