

Effects of Collaborative Training Using Virtual Co-embodiment on Motor Skill Learning

Daiki Kodama , Takato Mizuho , Yuji Hatada , Takuji Narumi , and Michitaka Hirose 

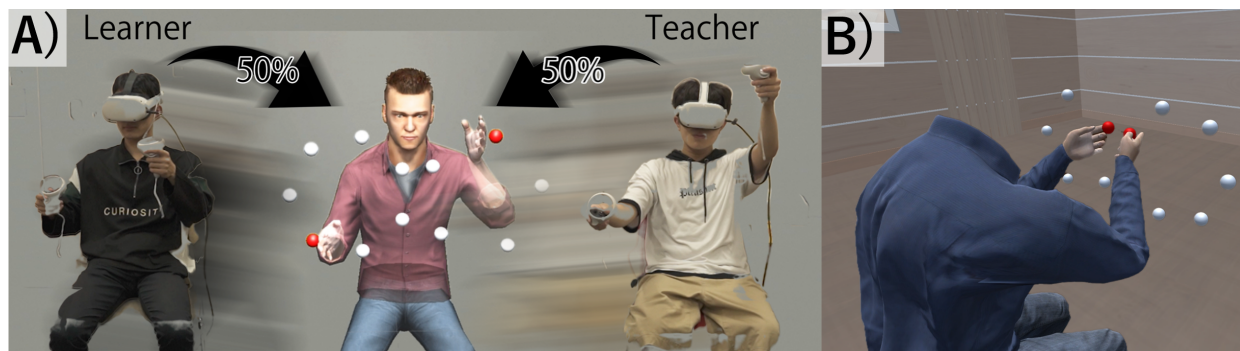


Fig. 1: **A)** Virtual co-embodiment: A virtual avatar is controlled based on the weighted average of two users' movements. This is the first study applying virtual co-embodiment to motor skill transfer. The current system is implemented following the work done by Fribourg et al [19]. **B)** Dual task: Participants draw a seven-pointed star with the right hand and a five-pointed star with the left. We found that learning dual task movements using virtual co-embodiment with a teacher enhances the retention of motor skills compared with sharing the teacher's first-person perspective and learning alone.

Abstract—Virtual reality (VR) is a promising tool for motor skill learning. Previous studies have indicated that observing and following a teacher's movements from a first-person perspective using VR facilitates motor skill learning. Conversely, it has also been pointed out that this learning method makes the learner so strongly aware of the need to follow that it weakens their sense of agency (SoA) for motor skills and prevents them from updating the body schema, thereby preventing long-term retention of motor skills. To address this problem, we propose applying “virtual co-embodiment” to motor skill learning. Virtual co-embodiment is a system in which a virtual avatar is controlled based on the weighted average of the movements of multiple entities. Because users in virtual co-embodiment overestimate their SoA, we hypothesized that learning using virtual co-embodiment with a teacher would improve motor skill retention. In this study, we focused on learning a dual task to evaluate the automation of movement, which is considered an essential element of motor skills. As a result, learning in virtual co-embodiment with the teacher improves motor skill learning efficiency compared with sharing the teacher's first-person perspective or learning alone.

Index Terms—Collaborative training, Virtual co-embodiment, Motor skill learning, Sense of agency



1 INTRODUCTION

The acquisition and long-term retention of motor skills play an essential role in various activities such as sports [28], industrial work [5], and nursing [6]. Typical motor skill learning involves observing the teacher's movements from a third-person perspective (3PP). Although 3PP observation has traditionally been employed as a method for learning motor skills, it can induce numerous errors in the cognitive process of the learner when they replicate movements observed in the teacher's body coordinates [53]. These errors prevent learners from understanding the correct movement and reduce the efficiency of motor skill learning [53].

To address this problem, many systems using virtual reality (VR) have been developed that allow learners to observe their teacher's movements from the teacher's first-person perspective (1PP) [34, 38, 69]. For instance, Yang et al. developed the “Just Follow Me” system [69], which superimposes a teacher's movements as a translucent avatar on the learner's body, and users observe it from 1PP in a virtual environment. Just Follow Me also allows learners to move their own bodies

by understanding and following the teacher's movements. They implemented the task of reproducing a specific trajectory by moving the hand and reported that the learner could reproduce the motions more accurately using Just Follow Me than with 3PP observation. However, it was also shown that although the learners could move accurately immediately after learning, they could not retain the motor skills learned. This is because the system does not enable the learner to understand why the teacher made a certain movement (motor intention [14]), whether it is from 1PP or 3PP, simply by observing and imitating it.

Updating the body schema plays a critical role in skill retention [55]. Body schema is assumed to be a common unconscious representation that supports all types of motor actions [9, 22]. The body schema is updated through body movements with a high sense of agency (SoA). SoA is the subjective feeling of initiating and controlling an action [7] and arises when the predicted outcome is perceived to be consistent with the actual outcome [11]. When learning requires a strong focus on following the body movements of others, SoA weakens because there is no need to predict the outcome of the action. If a strong SoA can be evoked for correct body movements, the acquisition of a correct body schema can be promoted and the efficiency of motor skill learning can be improved.

Meanwhile, a system called “virtual co-embodiment” [19] has been proposed. Virtual co-embodiment is a system in which a virtual avatar is controlled based on the weighted average of the movements of multiple users, and each user embodies that avatar. Through multiple experiments, Fribourg et al. reported that users overestimate their SoA when using virtual co-embodiment, and that the movements of two

- Daiki Kodama, Takato Mizuho, Yuji Hatada, Takuji Narumi, and Michitaka Hirose are with the University of Tokyo, Tokyo 113-8654, Japan. E-mail: {d_kodama | takato | hatada | narumi | hirose}@cyber.t.u-tokyo.ac.jp.

users tend to be aligned when the goal is shared, suggesting that there may be shared motor intention with its use [19]. These features may help address some of the problems with observation-based learning described earlier. Therefore, we propose a collaborative training method in which a teacher and a learner use virtual co-embodiment and motor skills are efficiently transferred from the former to the latter. Using this configuration, learners can experience body movements with a strong SoA, as if they themselves performed these movements, which is appropriately compensated by mixing the teacher's movements with their own movements. It is then expected that the transmission of motor intention from the teacher to the student will not only enable the student to demonstrate appropriate motor skills during co-embodiment but also update their body schema and retain motor skills after the training. To test the effectiveness of the proposed method, we investigated the extent to which learning using the proposed and existing methods is effective for motor learning through a dual-task paradigm.

The contributions of this study are as follows:

- It represents the first proposed application of virtual co-embodiment to motor skill transfer.
- It shows that learning using virtual co-embodiment with the teacher improves motor skill learning efficiency compared to sharing the teacher's IPP or learning alone.

2 RELATED WORK

2.1 Motor Skill Learning

Motor skill learning is the process of increasing the spatial and temporal accuracy of movements through practice [66, 67]. It consists of three phases [62] – cognitive, associative, and autonomous – with each phase being sequential and subsequent [2, 18, 59].

First, in the cognitive phase, learners strive to understand the correct movements, which contain implicit knowledge that is difficult to convey verbally [31]. Traditional IPP VR systems for motor skill learning have focused on this phase; for instance, for improving walking efficiency [46] and calligraphy tasks [58, 68, 69].

Next, following the cognitive phase, learners put their knowledge into practice during the associative phase. Whether learners can move correctly is irrelevant to knowing the correct movements [12]; thus, they must practice moving their bodies consciously and repeatedly.

Finally, in the autonomous phase, learners can gradually perform their skills unconsciously. After this phase, the performance becomes smooth, effortless, and fast in any environment [15, 33]. The unconscious body image acquired in this phase is called body schema [9, 22]. For instance, when attempting to grasp an object accurately, people calculate how much they should move their body parts based on the body schema [13, 56]. Additionally, body schema is known to be plastic in nature [10, 49]. When the body schema is appropriately updated, learners can unconsciously perform their skills [50]. For instance, Iriki et al. [35] trained macaque monkeys to retrieve distant objects using a rake. After training, the monkey's visual receptive fields of bimodal neurons were altered to include the entire length of the rake, implying that the body schema was updated to use the rake.

Body schema does not automatically get updated simply by transforming the body shape or grasping a tool in a virtual environment. The SoA plays an important role in updating the body schema [16]. The SoA represents the subjective feeling of initiating and controlling an action [7, 37]. In IPP VR, the learner does not feel a strong SoA for updating body schema because they focus on tracing the teacher's movement. As a result, they cannot similarly demonstrate the skill the next day [58, 68]. Compared to IPP VR systems, virtual co-embodiment can induce a stronger SoA for more appropriate body movements in which teacher and student movements are integrated. Therefore, we hypothesize that co-embodiment can realize motor skill learning as if the teacher's motor skills are transferred to the students, and the efficiency of motor skill learning is superior to that of traditional methods.

To measure whether the body schema is updated during this automation phase, the dual-task paradigm is commonly used [40]. Here, users simultaneously perform two simple tasks [54, 63]. It has been reported that participants can perform the dual task better after practice [17], and

high-skilled sports players can perform the dual task better than low-skilled ones because more skilled participants can focus on multiple objects with a lower cognitive load [20, 21].

2.2 Virtual Co-embodiment

Virtual co-embodiment is a system that enables a user to share a virtual avatar with another entity [19]. Virtual co-embodiment consists of two configuration methods: weighted-average-based and body-part-segmented based. In this section, we describe each of these methods.

2.2.1 Weighted-average-based Virtual Co-embodiment

Weighted-average-based virtual co-embodiment is a system in which a virtual avatar is controlled based on the weighted average of the movements of multiple entities. The percentage of each entity's movement contributing to the avatar is called the "weight." In weighted-average-based virtual co-embodiment, it is known that the user's sensation toward the avatar changes according to their weight [19, 27]. For example, it is known that SoA and sense of body ownership (SoBO) increase as the user's weight increases. Therefore, setting an appropriate weight is essential to ensure that SoA is required for motor skill learning. The strength of the SoA varies depending on the type of user movements [19]. For example, compared with free actions, actions with a fixed target position or trajectory tend to generate a stronger SoA because of the ease of movement prediction. Furthermore, they suggested that the interaction between the two users could lead to sharing of motor intention and motor synchronization in the weighted-average-based virtual co-embodiment [19]. This sharing motor intention could be applied to skill transfer. It is known that the avatar's hand movements become straighter and less jerky than the participant's hand movements in the weighted-average-based virtual co-embodiment with 50% weight [26]. In addition, the user can prioritize the movement of a co-embodied avatar rather than their own body. This knowledge indicates that the teacher may be able to move a co-embodied avatar so that it performs the skill correctly even when the learner's incorrect movement is reflected.

2.2.2 Body-part-segmented Virtual Co-embodiment

Body-part-segmented virtual co-embodiment is a system in which two users simultaneously manipulate different body parts. Hapuarachchi et al. [29, 30] developed this method in which two individuals controlled their left or right limbs. They found that embodiment toward the arm controlled by the partner was significantly higher when the participant dyads shared a common objective or when they were allowed to see their partner's goal, compared to when their partner's goal was unknown to them.

In body-part-segmented virtual co-embodiment, only pre-selected body parts are reflected in the teacher's movements, making it difficult for the learner to learn how to move other body parts. In addition, to properly transfer skills, it is necessary to appropriately determine which body parts of the co-embodied avatar are moved by the teacher and which by the students. However, the optimal method of mapping body parts is considered to vary from skill to skill, and no guidelines for designing such a mapping method have been investigated to date. Therefore, in this study, we adopted a weighted-average-based virtual co-embodiment because the teacher can move all the body parts of the co-embodied avatar to perform the skill correctly. Additionally, we adopted 50% as the weight commonly used in previous studies.

3 EXPERIMENT

This study was conducted to test whether the virtual co-embodiment improves motor skill learning efficiency. Participants learned the skill of performing two simple tasks simultaneously using three methods: Virtual Co-embodiment, Perspective Sharing, or Alone. The efficiency of motor skill learning was defined based on performance improvement from the baseline. This study was approved by the ethics committee of Graduate School of Information Science and Technology, the University of Tokyo (UT-IST-RE-220901-32).

Table 1: Defined terms for the calculation of the virtual co-embodiment system

Variable	Definition
ω	Control weight of the learner ($0 \leq \omega \leq 1$)
$x_{learner}$	Hand position of the learner
$x_{teacher}$	Hand position of the teacher
x_{shared}	Hand position of co-embodied avatar
$q_{learner}$	Hand quaternion of the learner
$q_{teacher}$	Hand quaternion of the teacher
q_{shared}	Hand quaternion of co-embodied avatar

3.1 Participants

Sixty-four participants were recruited for the experiment [53 males, 11 females, average age = 23.5 ± 3.4 (SD)]. All participants were unaware of the purpose of the experiment and had a normal or corrected-to-normal vision. Two participants had no previous experience with VR, 39 had limited previous experience with VR, and 23 were familiar with VR.

3.2 Virtual Co-Embodiment Platform

A virtual co-embodiment system was developed using Unity (version 2020.3.2f1). Our setup was based on two computers, two Meta Quest 2 head-mounted displays (HMDs), and two pairs of controllers to immerse the teacher and learner in a virtual environment. The participants were embodied in a virtual avatar from the 1PP. We used the male avatar (Male_Adult_08) for men and the female avatar (Female_Adult_05) for women from the Microsoft Rocketbox Avatar Library [25].

The movement of the co-embodied avatar is calculated as follows: Table 1 presents the variables used in the calculations. The co-embodied avatar's hand position (x_{shared}) can be calculated based on the weighted average of the learner's controller position ($x_{learner}$) and the teacher's controller position ($x_{teacher}$). Similarly, the co-embodied avatar's hand quaternion (q_{shared}) can be calculated based on the weighted average of the learner's controller quaternion ($q_{learner}$) and the teacher's controller quaternion ($q_{teacher}$). These can be described by the following equations:

$$x_{shared} = \omega x_{learner} + (1 - \omega) x_{teacher} \quad (0 \leq \omega \leq 1) \quad (1)$$

$$q_{shared} = \omega q_{learner} + (1 - \omega) q_{teacher} \quad (0 \leq \omega \leq 1) \quad (2)$$

The positions and postures calculated for each frame based on these equations were reflected in the co-embodied avatar using the Final IK Unity package¹. Final IK is a plug-in to compute full body motion based on inverse kinematics from the position and rotation of the head and hands.

In this study, the head was not co-embodied with the collaborative partner, based on Fribourg's virtual co-embodiment system [19]. A preliminary experiment ($N = 2$) indicated that the user became sick in VR when sharing the head positions because the user's vision moved regardless of their intention. Furthermore, in the preliminary experiment, the avatar's head was fixed such that it sometimes entered the field of view and distracted the user's immersion when they moved their head backward. Therefore, we made the avatar's head transparent to prevent it from entering the user's view.

3.3 Dual Task

In this study, we adopted a dual-task paradigm, which is a procedure that requires an individual to perform two tasks simultaneously, as a commonly used method to measure whether a body schema is updated [40]. The task included the simultaneous drawing of a seven-pointed star with the right hand and a five-pointed star with the left hand. Figure 2 shows a participant performing the dual task. Spheres were displayed at the vertices of the five and seven-pointed stars, and the participants drew figures by touching the vertices in turn. Figure 3

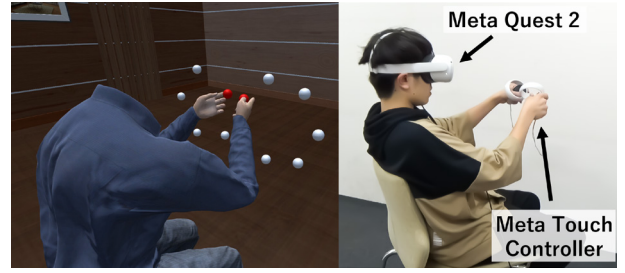


Fig. 2: A participant performing the dual task. The head of the avatar was kept transparent so that it did not interfere with the task.

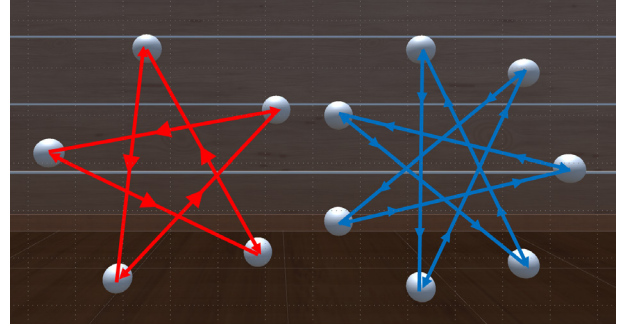


Fig. 3: Spheres were displayed at the position of the vertices of the five- and seven-pointed stars. The participants drew figures by touching the vertices in turn. The red and blue lines were not visible to the participants.

shows the hand movements to be made by the participant with lines and arrows. The red and blue lines in Fig. 3 are for illustrative purposes and were not actually visible to the participants. When the participants touched the vertex, vibration feedback was provided for 0.1 s. The color of the vertex changed to red when the participants touched the correct vertex and to yellow when they touched the wrong vertex, as shown in Fig. 4. The participants were informed of the current number of trials, number of times they touched the correct vertices, number of mistakes, remaining time, and current experimental phase (attempt or rest) from the board in front of them.

One trial of the dual task consisted of a 30 s task attempt and a 30 s rest phase. At the beginning of the task attempt phase, the color of one vertex in each figure changed to red, and participants started the task by touching the vertex. The vertex that the participants should have touched first was randomly selected. In the rest phase, all vertices became white, did not change color, and provided no vibration feedback when touched by the participants.

3.4 Design

The experiment was conducted using a mixed design. The independent variable between participants was the learning method with three levels:

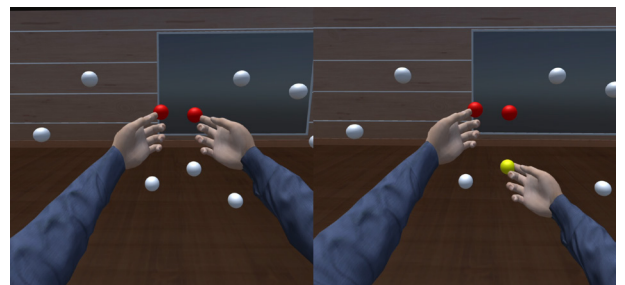


Fig. 4: Feedback on whether the participants touched the correct vertex. As shown on the left, the color of vertex changed to red when the participant touched the correct vertex. Conversely, as shown on the right, the color of vertex changed to yellow when the participants touched the wrong vertex. Vibration feedback was given to the participants for 0.1 s when they touched the vertex.

¹<https://assetstore.unity.com/packages/tools/animation/final-ik-14290>



Fig. 5: Perspective Sharing. The avatar in vintage red clothes represents the teacher's avatar. The participant performed the dual task by sharing the teacher's perspective, which superimposes the teacher's movements as a translucent avatar on the participants' IPP of their body in the learning phase.

Co-embodiment, Perspective Sharing, and Alone. Participants were randomly assigned in advance to three conditions (Co-embodiment: 19 males, 3 females, average age = 22.7 ± 2.1 (*SD*), Perspective Sharing: 17 males, 4 females, average age = 24.2 ± 4.0 (*SD*), Alone: 17 males, 4 females, average age = 23.5 ± 3.8 (*SD*)). The independent variable within the participants was the trial (11 levels: first, second, ..., tenth, *test*).

Co-embodiment: The participant performed the dual task using the virtual co-embodiment during the learning phase. The collaborative teacher of the virtual co-embodiment was a skilled experimenter. The weight ω (shown in Tab. 1) was set to 50%, as the teacher and learner could recognize each other's movement.

Perspective Sharing: The participant performed the dual task by sharing the teacher's perspective. Figure 5 shows the scene presented to both the teacher and the participants. The teacher's movements were superimposed as a translucent avatar on the participants' IPP of their bodies in the learning phase. The teacher was a skilled experimenter.

Alone: The participant performed the dual task alone in the learning phase.

The same experimenter (one of the authors) was in charge of the teacher's role for all participants in both the Co-embodiment and Perspective Sharing conditions. The author played the role of the teacher because the teacher had to become very proficient with virtual co-embodiment, perspective sharing, and dual task. These proficiencies were prioritized over a limitation that the author knew the hypothesis and would make it easier to get desirable results. In both Co-embodiment and Perspective Sharing, the teacher consciously tried to move toward the next vertex slightly earlier than the participants to guide the learner's movements. He practiced the dual task in advance and until he gained sufficient proficiency. The fact that no participant became more proficient than him in this experiment proves that was proficient enough. In addition, he was assisted by pink lines indicating the correct direction and vertex to move next to stabilize his performance. This pink line was not presented to the participants.

3.5 Procedure

Figure 6 summarizes this procedure. Participants were briefed and signed a consent form to participate in the experiment. They were briefed on the dual task. At this time, they were instructed to perform the task as fast as possible without touching the incorrect vertex. For Co-embodiment and Perspective Sharing, they were explained the system they would use. As a tutorial for the dual task, they wore HMDs and confirmed that the correct vertex would turn red with vibration feedback when they touched a vertex, and the wrong vertex would turn yellow with no vibration. In the case of Co-embodiment, participants experienced virtual co-embodiment while they experienced sharing the teacher's perspective in Perspective after the tutorial. The participants first performed the dual task once alone as a baseline. They

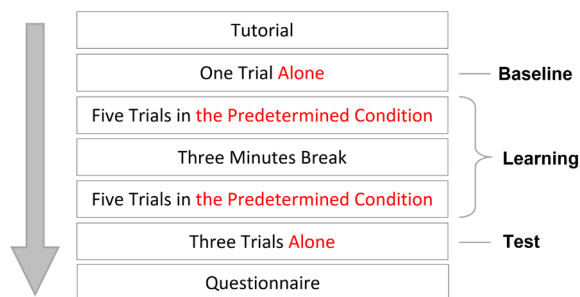


Fig. 6: The procedure of the experiment. One trial of the dual task consisted of a 30 s task attempt and a 30 s rest phase.

then performed the dual task five times as a learning phase under predetermined condition, that is, Co-embodiment, Perspective Sharing, or Alone. After removing the HMDs and resting for 3 min, they performed the dual task another 5 times under the same condition. Finally, the participants performed three trials of the dual task alone as part of the test. After they completed all trials, they removed HMDs, answered the virtual embodiment (VEQ) questionnaire (as described in the following section), and orally commented on the dual task and the system they used. The participants were asked to provide data regarding their age and experience with VR along with the VEQ. The total experimental time was approximately 45 minutes.

3.6 Measurement

3.6.1 Task performance

We compared the following measurements in the dual task between the learning method. We evaluated task performance as the number of times participants touched the correct vertex. First, we measured task performance in the tutorial and used it as the *baseline* for the participant. Second, for each trial, the *improvement* was defined as the task performance at that point minus the *baseline* and was used to measure the learning effect. Third, as a measure of learning retention, the *test score* was defined as the maximum *improvement* of the three test trials in the test phase in that the participants performed the task alone after the learning phase. The *test score* was used as an indicator of the efficiency of motor skill learning for the participant. Finally, as an indicator of the degree of residual learning outcomes, we defined the *performance drop* as how much performance is reduced when the assistance is terminated. *Performance drop* was calculated as the *improvement* in the last trial of the learning phase minus the *test score*.

3.6.2 Hand Distance

In addition, because it is known that interaction using virtual co-embodiment can influence the two users' movements [19, 26], we also recorded the trajectory of the hands. Hagiwara et al. showed that participants coordinated their movements with each other so that the co-embodied avatar moves in a more goal-oriented manner [26]. They showed that the positions of the two users' hands became further apart after the continuous use of the virtual co-embodiment. This movement coordination according to the situation may be useful when executing a movement in virtual co-embodiment, but may not be functional without partners. On the contrary, Hagiwara et al. have shown that such partner-aware movement changes appear using a simple reaching task, but the task used here involves more complex movement. Therefore, it is not obvious whether such partner-aware movement changes are observed in the present experimental setup. In Perspective-Sharing, it is expected that the teacher and learner's movements become closer as the learning progresses. Conversely, following previous research [26], the teacher and learner's movements are expected to become further apart under the virtual co-embodiment condition. As a result, it is possible that the learner may not be able to retain the assigned task after the assistance of virtual co-embodiment was terminated. Considering these points, the change in two users' movements may also affect the final learning effect. Therefore, we calculated the Euclidean distance of the

teacher's and learner's hands for each frame and defined *hand distance* as the average distance within a trial. We evaluated two main things: (a) the change in *hand distance* and (b) a relationship between the *hand distance* and the *performance drop*.

3.6.3 Sense of Embodiment

We measured indices that described the degree of embodiment in the virtual avatar: SoA, the SoBO, and Change. Along with SoA, SoBO is a component of the sense of embodiment that emerges when the avatar's properties are processed as if they are the properties of one's own biological body [43]. Change refers to the change in the perceived body schema due to stimulation [57]. We used a virtual embodiment questionnaire [57] (VEQ) to assess SoA, SoBO, and Change to confirm that participants embodied in their avatars. The VEQ is a commonly used questionnaire that can be applied in various VR experiments [57].

3.7 Hypothesis

The learner is expected to be able to improve their movements more efficiently during the learning phase because the learner receives feedback on the teacher's correct movements in the Co-embodiment and Perspective Sharing

H_1 : The *improvement* in the Co-embodiment and Perspective Sharing conditions is higher than that in the Alone condition during the learning phase.

In the Perspective Sharing condition, participants are so focused on following the teacher's movements that they do not have room to understand the teacher's motor intentions, which is expected to result in a significant *performance drop*. On the other hand, participants in the Co-embodiment condition are expected to have a stronger SoA to the teacher's movements and could learn the teacher's motor intentions, thus relatively reducing the *performance drop*.

H_{2-1} : The *performance drop* in the Perspective Sharing condition is significantly greater than that in the Alone condition.

H_{2-2} : The *performance drop* in the Co-embodiment condition is smaller than that in the Perspective Sharing condition.

Finally, the *test score* measured without teacher assistance is expected to be smaller because participants in the Perspective Sharing condition do not have the room to understand the teacher's motor intentions compared to participants in the Co-embodiment condition.

H_3 : The *test score* in the Co-embodiment condition is higher than those in the Perspective Sharing and Alone conditions.

4 RESULT

The data from one participant who could not complete the experiment due to a system error was removed. Data from the remaining 63 participants were analyzed. The significance level was set at $p < 0.05$.

4.1 Baseline

To ensure that there was no bias in the assignment to learning method, we verified whether there were differences in the *baseline*. Figure 7a shows the *baseline* for each condition. One-way ANOVA with the between-subjects factor learning method (three levels: Co-embodiment, Perspective Sharing, and Alone) was performed for the *baseline* because the normality assumption (Shapiro-Wilk's normality test) was not violated. The results of one-way ANOVA showed no significant differences between baselines ($F(2,60) = 1.48, p = .24, \eta_p^2 = 0.05$).

4.2 Skill Learning

4.2.1 Improvement

Figure 8 shows a line chart of the *improvement* and the observed *testing score*. Because the normality assumption (Shapiro-Wilk's normality test) was not violated, two-way ANOVAs with the between-subjects factor learning method (three levels: Co-embodiment, Perspective Sharing, and Alone) and within-subjects factor trials (ten levels: first, second, ..., tenth) were performed for *improvement*. The two-way ANOVAs revealed a significant main effect of the learning method ($F(2, 60) = 94.04, p < .001, \eta_p^2 = 0.76$) and of trial ($F(9, 18) = 95.96,$

$p < .001, \eta_p^2 = 0.62$). Because the two-way ANOVAs also exhibited a significant interaction effect between the learning method and trial ($F(18, 540) = 6.93, p < .001, \eta_p^2 = 0.19$), a post-hoc analysis, the Welch's t-test adjusting p-value using the Bonferroni's method, was performed.

First, we tested whether participants improved their skills in the dual task during the learning phase. The first trial's *improvement* was compared with that of the tenth trial in each condition. As a result, the first trial's *improvement* was found to be significantly lower than that of the tenth trials for each learning method ($p < .01$ in all conditions). The result indicated that participants may have improved their skills in the dual task during the learning phase in all learning methods.

Second, we tested which conditions participants became most proficient in the learning phase. Post-hoc analysis between conditions was conducted on all learning trials. The results indicated that, for all trials, the *improvement* in the Co-embodiment condition was significantly higher than that in the Perspective Sharing ($p < .001$ in all learning trials) and Alone conditions ($p < .001$ in all learning trials). In addition, for all trials, the *improvement* in the Perspective Sharing condition was significantly higher than that in the Alone condition ($p < .01$ in all learning trials). As the ANOVA indicated that the strong effects originated from the trial factor, a linear regression analysis was conducted across trials for each method to further characterize the relationship between the *improvement* and trial. The regression equations were as follows:

Co-embodiment: $y = 1.98x + 14.13$ ($R^2 = 0.44$),

Perspective Sharing: $y = 1.14x + 7.50$ ($R^2 = 0.31$),

Alone: $y = 0.86x + 2.34$ ($R^2 = 0.25$).

The regression equations exhibited positive linear correlations between *improvement* and trial in each learning method. We then calculated the slope of *improvement* over trial for each participant. We conducted one-way ANOVA to compare the regression slopes among learning methods because the normality and homogeneity of variance assumptions (Shapiro-Wilk's normality test and Levene test) were not violated for the slope. The Welch's t-test adjusting p-value using the Shaffer's method was performed because the result showed the significant main effect of the learning method ($F(2, 60) = 24.45, p < .001, \eta_p^2 = 0.45$). The result showed that the slope of *improvement* under the Co-embodiment condition was significantly higher than that under the other two conditions (Perspective Sharing: $t(60) = 6.72, p < .001$, Alone: $t(60) = 5.02, p < .001$). Furthermore, the result reported that the slope of *improvement* under the Perspective Sharing condition was marginally higher than that under the Alone condition ($t(60) = 1.70, p = .094$). These results indicated that the slope of *improvement* during learning phase was significantly higher in the order of the Co-embodiment, the Perspective Sharing, and the Alone conditions. In addition, the *improvement* during the learning phase was significantly higher in the order of the Co-embodiment, the Perspective Sharing, and the Alone conditions. These results support **H1**.

4.2.2 Hand Distance

Figure 7b shows the number of trials and the hand distance from the teacher in the Co-embodiment and the Perspective Sharing conditions. Because the normality assumption (Shapiro-Wilk's normality test) was not violated, two-way ANOVAs with the between-subjects factor learning method (two levels: Co-embodiment and Perspective Sharing) and within-subjects factor trials (10 levels: first, second, ..., tenth) were performed for the *hand distance*. The two-way ANOVAs revealed a significant main effect of the learning method ($F(1, 40) = 37.50, p < .001, \eta_p^2 = 0.48$) and trial ($F(9, 9) = 4.04, p < .001, \eta_p^2 = 0.09$). Because the two-way ANOVAs also exhibited a significant interaction effect between the learning method and trial ($F(9, 360) = 11.23, p < .001, \eta_p^2 = 0.22$), a post-hoc analysis, the Welch's t-test adjusting p-value using Bonferroni's method was performed. First, *hand distance* in the first trial was compared with that in the tenth trial for each learning method. As a result, the *hand distance* in the first trial was found to be significantly lower than that in the tenth trial in the Co-embodiment condition ($p < .001$). Conversely, the *hand distance* in

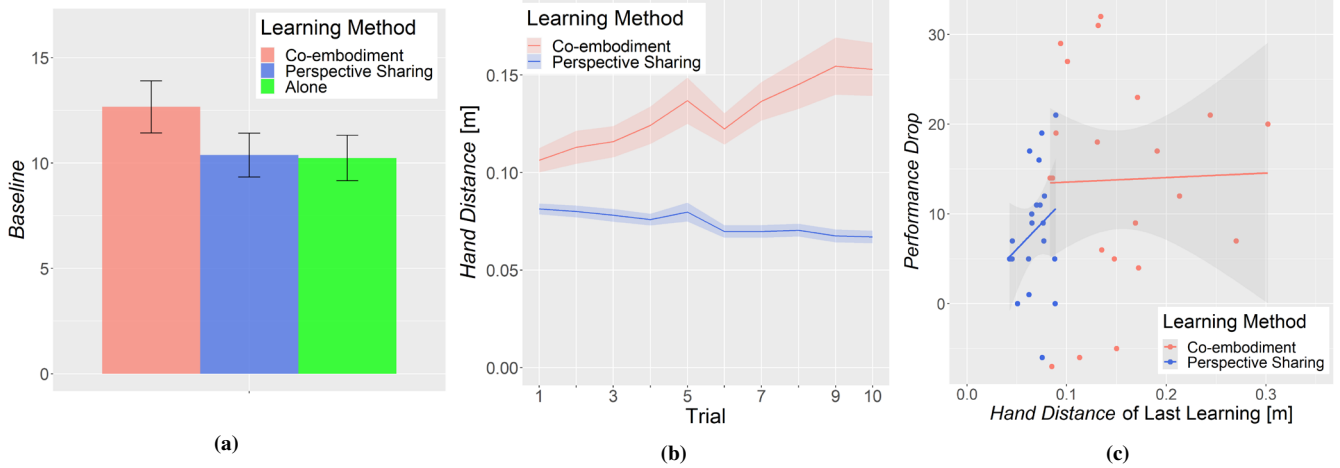


Fig. 7: (a) : The baseline in each condition. No significant differences were found among the three conditions. (b) : Graph of the number of trials and the hand distance during the learning phase. Error bands are standard errors. The hand distance was significantly greater in the Co-embodiment condition than in the Perspective Sharing condition. In the Co-embodiment condition, the hand distance tended to increase with the number of trials. On the contrary, in the Perspective Sharing condition, the hand distance tended to decrease with the number of trials. (c) : Point plot of the mean hand distance from the teacher at the last learning trial and performance drop in the Co-embodiment and Perspective Sharing conditions. No correlation was found in both conditions.

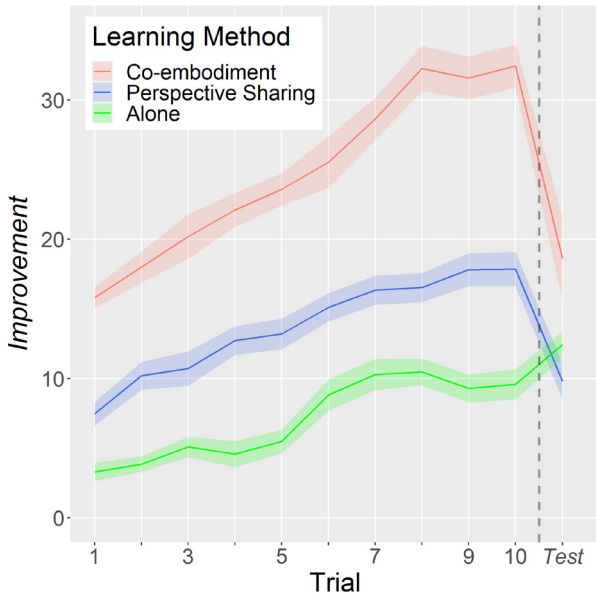


Fig. 8: Line chart of the improvement. Error bands are standard errors. Trials 1–10 show the improvement during the learning phase. Test shows the maximum improvement of the three test trials, where the participants performed the task alone. The improvement during the learning phase was higher in the order of the Co-embodiment, the Perspective Sharing, and the Alone conditions. The performance drop was significantly greater in the order of the Co-embodiment, Perspective Sharing, and Alone conditions. The test score in the Co-embodiment condition is higher than that in the Perspective Sharing or Alone conditions.

the first trial is significantly higher than that in the tenth trial in the Perspective Sharing condition ($p < .001$). The ANOVA indicated a strong relationship between the number of trials and hand distance in both learning methods. Therefore, a linear regression analysis of the number of trials and hand distance for each learning method was conducted to further characterize these relationships. The regression equations are as follows:

Co-embodiment: $y = 0.0053x + 0.102$ ($R^2 = 0.094$),

Perspective Sharing: $y = -0.0017x + 0.083$ ($R^2 = 0.096$).

The regression equations exhibited positive linear correlations between hand distance and trial in the Co-embodiment condition. In contrast,

the regression equations exhibited negative linear correlations between hand distance and trial in the Perspective Sharing condition. We computed Pearson's product-moment coefficient for each condition to determine whether the computed slopes differed significantly from 0. As a result, we found that the two variables were strongly correlated (Co-embodiment: $Pearson's r(208) = 0.31$, $p < .0001$, Perspective Sharing: $Pearson's r(208) = -0.31$, $p < .0001$). These results suggest that the distance between the participant's and teacher's hands may have gradually increased in the Co-embodiment condition. In contrast, the distance between the participant's and teacher's hands may have gradually decreased in the Perspective Sharing condition.

4.3 Skill Retention

4.3.1 Performance Drop

Because the normality assumption (Shapiro–Wilk's normality test) was not violated, one-way ANOVA with the between-subjects factor learning method (three levels: Co-embodiment, Perspective Sharing, and Alone) was performed for performance drop. Because the result showed a significant main effect of learning method ($F(2, 60) = 23.85$, $p < .001$, $\eta_p^2 = 0.44$), a post-hoc analysis, the Welch's t-test adjusting p-value using the Shaffer's method, was performed. The result indicated that performance drop was significantly greater in the order of the Co-embodiment, the Perspective Sharing, and the Alone conditions ($p < .0001$, between all possible pairs). This result supports **H2_1**, but not **H2_2**.

4.3.2 Test Score

Because of concerns about outliers from the histogram shape, we conducted the Grubbs test. The analysis did not detect any outliers. Because the normality assumption (Shapiro–Wilk's normality test) was not violated, one-way ANOVA with the between-subjects factor learning method (three levels: Co-embodiment, Perspective Sharing, and Alone) was performed for the test score. Because one-way ANOVA revealed a significant main effect of the learning method ($F(2, 60) = 5.33$, $p < .01$, $\eta_p^2 = 0.15$), a post-hoc analysis, the Welch's t-test adjusting p-value using the Shaffer's method, was performed. The result indicated that, the test score in the Co-embodiment condition was significantly higher than that in the Perspective Sharing ($t(60) = 3.18$, $p < .05$) and Alone conditions ($t(60) = 2.23$, $p < .05$). There was no significant difference between the Perspective Sharing and Alone conditions ($t(60) = 2.61$, $p = .35$). This result supports **H3**.

To investigate the possibility that the participants' proficiency may have influenced their test score, a plot of the correlation between base-

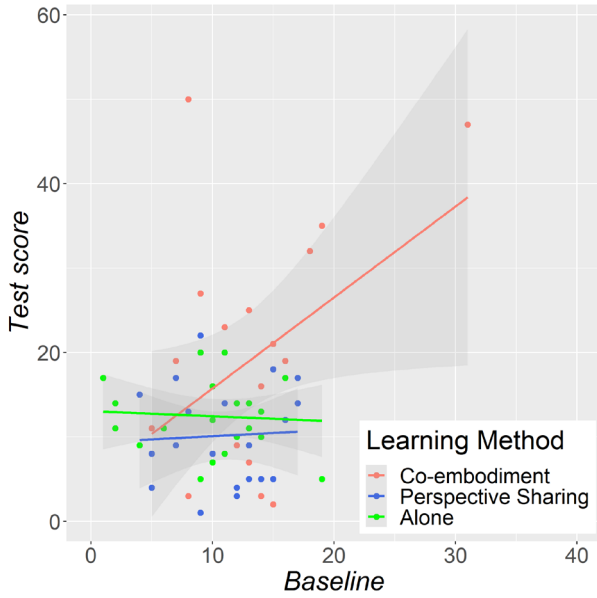


Fig. 9: Point plot of the *baseline* and the *test score*. No correlation was observed in the Perspective Sharing and Alone conditions. On the contrary, a positive correlation was observed in the Co-embodiment condition: the higher the baseline, the higher the learning efficiency.

line and *test score* in all learning methods is shown in Fig. 9. A regression analysis was performed for all learning methods.

Co-embodiment: $y = 1.08x + 4.97$ ($R^2 = 0.20$),

Perspective Sharing: $y = 0.23x + 7.40$ ($R^2 = 0.034$),

Alone: $y = -0.06x + 13.0$ ($R^2 = 0.004$).

We computed Pearson's product-moment coefficient for each learning method to determine whether the computed slopes differed significantly from 0. As a result, there was a significant positive correlation under the Co-embodiment condition: the higher the baseline, the more efficient learning (Pearson's $r(19) = 0.45$, $p < .05$). On the other hand, there was no significant correlation under Perspective Sharing (Pearson's $r(19) = 0.19$, $p = .42$) and Alone conditions (Pearson's $r(19) = -0.07$, $p = .77$).

4.3.3 Correlation between Hand Distance and Performance Drop

Figure 7c shows a point plot of the mean *hand distance* at the last learning trials and the *performance drop* in the Co-embodiment and Perspective Sharing conditions. A linear regression analysis was performed, resulting in the following regression equation:

Co-embodiment: $y = 5.0x + 13.0$ ($R^2 = 0.012$),

Perspective Sharing: $y = 116.9x + 0.27$ ($R^2 = 0.014$).

We computed Pearson's product-moment coefficient for each learning method to determine whether the computed slopes differed significantly from 0. As a result, no correlation was confirmed in both learning methods (Co-embodiment: Pearson's $r(19) = 0.03$, $p = .90$, Perspective Sharing; Pearson's $r(19) = 0.25$, $p = .26$).

4.4 Sense of Embodiment

Figure 10 shows the scores of the sense of embodiment. The Friedman test for the SoA, SoBO, and Change (Fig. 10a, Fig. 10b, and Fig. 10c) were conducted with the between-subjects factor learning methods (three levels: Co-embodiment, Perspective Sharing and Alone). The results showed no significant difference for each measurement (SoA: $\chi^2 = 1.81$, $p = .40$, SoBO: $\chi^2 = 0.22$, $p = .89$, Change: $\chi^2 = 0.04$, $p = .98$).

5 DISCUSSION

5.1 Skill Learning

5.1.1 Higher Improvement under Conditions with the Teacher's Assistance

The *improvement* in the Co-embodiment and Perspective Sharing conditions during the learning phase was significantly higher than in the Alone condition (Sec. 4.2.1). The results were consistent with the previous work reporting that dyadic interaction allows the user to estimate the motion intention of the partner based on feedback and modify their motion [4, 24, 51]. In the Perspective Sharing condition, the participants could improve their performance during the learning phase by observing the translucent teacher avatar's movements and modifying their movements. In contrast, in the Co-embodiment condition, performance was improved during the learning phase simply because the teacher could correct the avatar's movements directly. Thus, the *improvement* in the Co-embodiment condition did not purely represent learners' improvement.

5.1.2 Various Trend of Hand Distance between learning methods

Figure 7b shows the *hand distance* had gradually increased in the Co-embodiment condition (Sec. 4.2.2). This result indicates that the user's behavior changed due to the interaction using virtual co-embodiment in the more complex task of this study, as seen in simple tasks such as reaching in previous studies [19, 26]. On the contrary, the distance between the participant's and teacher's hands may have gradually decreased in the Perspective Sharing condition. This difference in correlation may indicate a difference in learning methods between the Co-embodiment and Perspective Sharing conditions. In the Perspective Sharing condition, participants might have learned by feedback error learning [41]. Feedback error learning is the process of learning by diminishing the error between the intended motor prediction and the perceived result [36, 42]. In the Perspective Sharing condition, feedback error learning that diminished the error with the teacher is likely to have occurred because the learner eliminated the difference between their movement and that of the teacher. On the other hand, in the Co-embodiment condition, feedback error learning is unlikely to occur because the hand distance from the teacher gradually increases. In the Co-embodiment condition, the learning mechanism may differ from the feedback error learning and does not necessarily correspond to making their real body movements closer to the teacher. This possibility was supported by the no correlation between the *hand distance* of the last learning trial and *performance drop* in the Co-embodiment condition (Shown in Fig. 7c). The participant's comment also supports this possibility, "In the latter half, I no longer felt that I was assisted by the teacher and was moving as my own body." This participant may have become less conscious of their own bodies as their learning progressed. It is also possible that the learner was distracted and slacked off by the reflection of the teacher's actions. In this case, further improvement in efficiency can be expected by motivating the learner. Either way, it is necessary to examine the mechanism of motor skill transfer through virtual co-embodiment.

5.2 Skill Retention

5.2.1 Significant Performance Drop under Conditions with the Teacher's Assistance

The *performance drop* in the Perspective Sharing condition was significantly greater than that in the Alone condition (Sec. 4.3.1). The results were consistent with the previous work reporting that participants depended so much on force feedback during learning and could not perform well on their own during testing in motor skill learning with haptic feedback [39]. Several participants support this by mentioning that the loss of the teacher's assistance made them unsure of what to do during the test. Certain participants also stated that the teacher's translucent avatar sometimes moved before they began to think about how to move. From these comments, it can be inferred that participants were so focused on following the teacher's movements that they could not learn actively.

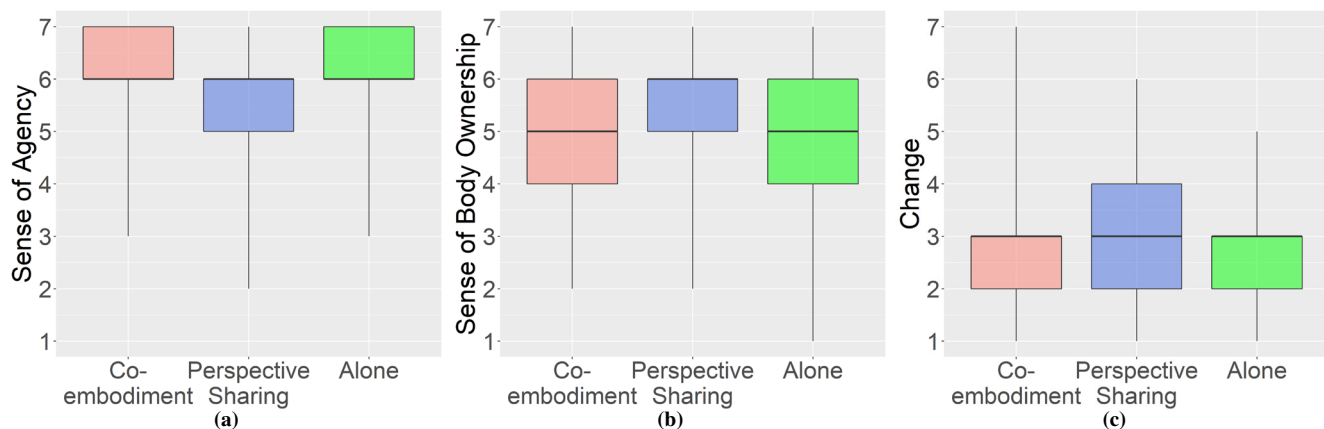


Fig. 10: Box-and-whisker diagram of sense of embodiment. No significant differences were found among the three learning methods. (a) Box-and-whisker diagram of SoA. SoA refers to the perception of control over a virtual body and, through that, control over events in the environment [57]. (b) Box-and-whisker diagram of SoBO. SoBO refers to the perception of perceiving a virtual body as the own, as the source of all sensations [57]. (c) Box-and-whisker diagram of Change. Change refers to the change in the perceived body schema due to stimulation [57].

The *performance drop* in the Co-embodiment condition was significantly larger than that in the Perspective Sharing condition (Sec. 4.3.1). In addition, the *performance drop* in the Co-embodiment condition was significantly larger than that in the Alone condition. The most significant *performance drop* in the Co-embodiment condition may be due to a significant *improvement* by the assistance. It is known that excessive system assistance tends to improve performance during learning but is also known to be prone to cause performance drop [39]. System assistance might have been most significant in the Co-embodiment condition because the *improvement* in the learning phase was most outstanding in this condition. Enormous system assistance in Co-embodiment condition may have caused a most significant *performance drop*. As described in the following section, the effectiveness of the proposed method is supported by the *test score*.

5.2.2 Higher Test Score in the Co-embodiment condition

The *test score* in the Co-embodiment condition was significantly higher than that in the Perspective Sharing and Alone conditions (Sec. 4.3.2). During the learning phase, as expected, *improvement* was significantly higher in the Co-embodiment and Perspective Sharing conditions than in the Alone condition. The highest test scores were obtained in the co-embodiment condition, indicating that the virtual co-embodiment with the teacher allows efficient transfer of motor skills. On the other hand, contrary to the expectation, the *performance drop* in the Co-embodiment condition was significantly greater than in the Perspective Sharing and Alone conditions. Nevertheless, **H3** was supported. This result may be because the effect of *improvement* in the Co-embodiment condition during the learning phase was too significant to counteract the effect of *performance drop*.

The relationship between weight and proficiency of the learner in virtual co-embodiment can be found in the scatter plots (Fig. 9) of *baseline* and *test score* (Sec. 4.3.2). From this figure, it can be inferred that participants with higher *baselines* acquired motor skills with higher learning efficiency when using Co-embodiment. This trend is not observed in other conditions. Many participants in the Co-embodiment condition commented in support of this relationship. For example, one participant stated, “I felt that I could perform better by doing the task several times by myself or with high weights at first. After I become a little proficient at the task, I will be able to understand the motor intent of the teacher and acquire the motor skill more easily using virtual co-embodiment.” The virtual co-embodiment system implemented in this study was considered to be particularly effective for users with a certain level of proficiency. This is possible because those with a certain level of proficiency already understood the basic movements and were then able to improve and automate the skills while understanding their partner’s motor intentions. The points to

be learned may change depending on the level of proficiency prior to practice. And the appropriate weights may change depending on what is to be learned. We used a virtual co-embodiment of 50% weight in this study, but different weights may be effective for learners with different proficiency levels. In addition, dynamical weight control can be used to improve the efficiency of motor skill transfer because it can affect the perceived SoA [45]. It is worthwhile not only to clarify the effect of fixed weights but also to examine whether dynamically changing weight based on the learner’s proficiency level or motion pattern can efficiently convey the teacher’s motor intention.

5.3 Sense of Embodiment

No significant differences among all learning methods were observed in the sense of embodiment (Sec. 4.4). We could not confirm the difference of whether the participants predicted motor intention actively from the SoA score in this experiment. Fribourg et al. showed that the sense of embodiment is lower with virtual co-embodiment of 50% weight than with 100% (when the avatar is fully operated by the participant) [19]. Nevertheless, in the present study, there was no difference in the embodiment scores between the co-embodiment condition with the avatar that reflects only 50% of the participants’ movements and the perspective-sharing condition with the avatar that directly reflects the participants’ movements, and a strong SoA was evoked in both conditions. One possible reason for this may be that the experience of success in the difficult task generated an even greater SoA than in Fribourg’s experiment. Wen et al. showed that success in a task enhances the sense of agency [64, 65]. Such an effect was likely observed in the co-embodiment with the teacher, where a high learning efficiency was observed. Another reason could be that in the perspective-sharing condition, users were so focused on following the teacher’s movements that their sense of agency was reduced.

As a limitation, note that the participants evaluated SoA by recalling the learning trial after they had performed the test trial alone (Fig. 6). The reason for the questionnaire timing was that our priority was seeing the skill retention immediately after the practice. They might overestimate the SoA because they experienced complete reflection of their movement in their avatar during the test.

6 FUTURE WORK AND LIMITATIONS

The results demonstrated that the proposed method could enhance the efficiency of motor skill learning. However, there were some limitations to the experimental method. There is a gap between the dual task and any other daily motor skills. We need to verify the proposed method’s applicability to different kinds of tasks. The dual task is a commonly used method to measure whether the body schema is updated [40]. The dual task used in this study is to perform different movements

with the left and right arms in parallel. Such movements are used in piano playing [1, 52] and drumming [44], and can be said to be an extraction of one of the characteristics commonly found in complex body movements. In the future, it will be necessary to verify the applicability of the proposed method to more practical skills such as piano and drums. Furthermore, it is also necessary to verify the proposed method's applicability to different tasks, such as tracing task, which requires reflexes. There is also a limitation in the design of the task. In this study, the number of learning sessions was set to 10. Figure 8 shows that learning did not completely saturate with the 10 learning sessions. This is also supported by the many participants' comments, such that their performance could have improved with more practice. To verify the relationship between proficiency and skill learning methods, the efficiency of skill learning after a more extended study period will need to be examined. In addition, in this study, the test trial was conducted only immediately after practice, and long-term skill retention was not investigated. In the future, it will be essential to investigate whether the learner can demonstrate the skill after a long retention interval.

How the teacher indicates movement to the learner might have also affected the efficiency of motor skill learning in virtual co-embodiment. In this study, the experimenter played the teacher's role. The experimenter tried to prioritize the movement of the co-embodied avatar rather than their own body and performed the movements slightly faster than the learner based on the learner's skill level. We observed the participant who felt that the teacher's approach worked well as well as that who felt that did not work well. The participant who felt that it worked well mentioned: "Because the teacher moved a little faster than me, there is little difference in movement between myself and co-embodied avatars. I learned the motor skill efficiently because I could feel that the co-embodied avatar's body was my own body." In contrast, the participant who did not feel it worked well mentioned: "I concentrated on thinking proactively and moving ahead of the teacher to acquire the skill. However, at the beginning of the test, I realized that the teacher had given me the correct answer a little faster than my movement, and I was just following it. Therefore, I could not demonstrate the skill without the teacher's assistance." The difference between these opinions may be whether they think of the co-embodied avatar as their self-body. The teacher might have to adjust how to move to make the learner recognize the co-embodied avatar as their own body. It is worth analyzing what kind of movement should be presented to what kind of learner, using the learner's personality traits such as locus of control (i.e., whether they believe that events in their life derive primarily from their own actions [48]). In addition, the present experiment did not examine the impact of collaborative training using virtual co-embodiment on the teacher's skill. As stated in Sec. 2.2.1, co-embodiment with students is not expected to negatively affect teachers, but rather a positive effect. On the contrary, it cannot be denied that prolonged exposure to the incorrect body movements of students may degrade the skills of teachers. It will be necessary to clarify the effect on the teacher side in future studies.

Further improvement in motor skill learning efficiency can be achieved by considering which characteristics of learners affect learning using virtual co-embodiment. For example, previous learning experience may affect efficiency. One participant commented: "It was easy for me to learn motor skills using virtual co-embodiment because my hand moving on its own was similar to the feeling I had when I was learning calligraphy with my teacher." Another participant commented: "I can learn the skill well using virtual co-embodiment because the feeling was similar to the feeling I had when I practiced on the piano. When I practiced one hand on the piano, the teacher took charge of the other hand, similar to virtual co-embodiment." It is possible that the learner who has had similar experiences could learn efficiently using virtual co-embodiment. It would be worthwhile to obtain personal traits and experiences and examine their relationship with motor skill learning.

Further analysis could be done by measuring movement dynamics (e.g., velocity and acceleration) and movement variability. Because the velocity of the learner will become close to that of the teacher as

learning proceeds [8], it might be essential to see how the learners' movement velocity responded to the teachers' movements as a function of training. Furthermore, movement variability could also be used as an indicator of proficiency because would become less variable as the participants develop proficiency in performing the task [47]. In this study, we have already analyzed a large number of measurements, and focused on *hand distance* and *improvement* because previous studies of virtual co-embodiment [19, 26] have dealt with hand distance and performance as a direct measurement motor skill learning efficiency. Further analyses would help us gain a deeper understanding of how different training modes affect learning efficacy.

Our result does not specify the mechanism by which motor intention is conveyed using virtual co-embodiment. It is worth verifying the mechanism using other evaluation measurements. For example, *we-mode* is a cognitive mode in which motor intention is communicated between the two [23]. In *we-mode*, interacting agents share their minds by representing their contributions to the joint action as contributions to something they will pursue together as a "we" [23]. Establishing *we-mode* makes it easier to acquire the interaction partner's point of view and understand the motor intention of the partner both potentially and automatically [32]. It is also known that the movements of individuals and brain waves tend to synchronize in *we-mode* [70]. Whether *we-mode* is established can be examined using button-pushing tasks such as the Simon task [60, 61] or the Flanker task [3] as well as brain wave measurement [70]. We may verify how motor intention transfer occurs by verifying whether *we-mode* is established when using virtual co-embodiment.

7 CONCLUSION

This is the first study to apply virtual co-embodiment to motor skill transfer. We verified the hypothesis that virtual co-embodiment with the teacher can help learners feel a strong SoA toward the correct movements taught to them, enabling the update of body schema and efficient skill transfer. We experimentally compared the Co-embodiment, Perspective Sharing, and Alone learning methods (N = 63). In Co-embodiment condition, the participants learned using virtual co-embodiment in which the teacher's movements also contributed to the participant's avatar. In the Perspective Sharing condition, the participants observed and followed the teacher's translucent avatar's movements from a 1PP. Finally, in the Alone condition, the participants learned alone. In all learning methods, participants first performed the dual task trial once alone as a baseline. The participants practiced the task 10 times under each condition with breaks in between. For all learning methods, participants performed the trials three times alone at the end as a test. The efficiency of motor skill learning was compared based on the differences between achievement in baseline and test trials. A comparison of 3 learning conditions showed that the Co-embodiment condition was significantly more efficient in motor skill learning than the other two conditions. This result supports the effectiveness of the proposed method. Furthermore, it is possible that the higher the performance on the baseline measurement, the more a learner is proficient when using the virtual co-embodiment. In the future, we plan to further improve the efficiency of skill transfer by dynamically controlling the weight in virtual co-embodiment according to the learner's proficiency level.

ACKNOWLEDGMENTS

This work was partially supported by Grant-in-Aid for Scientific Research (S) (19H05661) and JST Moonshot Research & Development Program (JPMJMS2013).

REFERENCES

- [1] D. A. Allport, B. Antonis, and P. Reynolds. On the division of attention: A disproof of the single channel hypothesis. *Quarterly journal of experimental psychology*, 24(2):225–235, 1972. doi: 10.1080/00335572430001029
- [2] J. R. Anderson. Acquisition of cognitive skill. *Psychological review*, 89(4):369, 1982. doi: 10.1146/annurev.psych.47.1.513 2

- [3] S. Atmaca, N. Sebanz, and G. Knoblich. The joint flanker effect: Sharing tasks with real and imagined co-actors. *Experimental brain research*, 211(3):371–385, 2011. doi: 10.1007/s00221-011-2709-9 9
- [4] E. J. Avila Mireles, J. Zenzeri, V. Squeri, P. Morasso, and D. De Santis. Skill learning and skill transfer mediated by cooperative haptic interaction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(7):832–843, 2017. doi: 10.1109/TNSRE.2017.2700839 7
- [5] A. P. Bartel and N. Sicherman. Technological change and the skill acquisition of young workers. *Journal of Labor Economics*, 16(4):718–755, 1998. doi: 10.1086/209904 1
- [6] I. T. Bjark. Changing conceptions of practical skill and skill acquisition in nursing education. *Nursing Inquiry*, 4(3):184–195, 1997. doi: 10.1111/j.1440-1800.1997.tb00098.x 1
- [7] N. Braun, S. Debener, N. Spsychala, E. Bongartz, P. Sörös, H. H. O. Müller, and A. Philipsen. The senses of agency and ownership: A review. *Frontiers in Psychology*, 9:535, 2018. doi: 10.3389/fpsyg.2018.00535 1, 2
- [8] A. W. Burton and D. E. Miller. *Movement skill assessment*. Human Kinetics, 1998. 9
- [9] L. Cardinali, C. Brozzoli, and A. Farne. Peripersonal space and body schema: two labels for the same concept? *Brain topography*, 21(3):252–260, 2009. doi: 10.1007/s10548-009-0092-7 1, 2
- [10] L. Cardinali, C. Brozzoli, J. Luaute, A. C. Roy, and A. Farnè. Proprioception is necessary for body schema plasticity: evidence from a deafferented patient. *Frontiers in Human Neuroscience*, 10:272, 2016. doi: 10.3389/fnhum.2016.00272 2
- [11] G. Carruthers. The case for the comparator model as an explanation of the sense of agency and its breakdowns. *Consciousness and cognition*, 21(1):30–45, 2012. doi: 10.1016/j.concog.2010.08.005 1
- [12] J. H. Caurraugh, M. Martin, and K. K. Martin. Modeling surgical expertise for motor skill acquisition. *The American journal of surgery*, 177(4):331–336, 1999. doi: 10.1016/S0002-9610(99)00057-4 2
- [13] F. De Vignemont. Body schema and body image—pros and cons. *Neuropsychologia*, 48(3):669–680, 2010. doi: 10.1016/j.neuropsychologia.2009.09.022 2
- [14] M. Desmurget and A. Sirigu. A parietal-premotor network for movement intention and motor awareness. *Trends in cognitive sciences*, 13(10):411–419, 2009. doi: 10.1016/j.tics.2009.08.001 1
- [15] S. T. Dumais. The development of automatism. *Cognitive Skills and Their Acquisition, Hillsdale, NJ: Erlbaum*, pp. 111–140, 1981. 2
- [16] M. D’Angelo, G. di Pellegrino, S. Seriani, P. Gallina, and F. Frassinetti. The sense of agency shapes body schema and peripersonal space. *Scientific reports*, 8(1):1–11, 2018. doi: 10.1038/s41598-018-32238-z 2
- [17] H. E. Ewolds, L. Bröker, R. F. De Oliveira, M. Raab, and S. Künnzell. Implicit and explicit knowledge both improve dual task performance in a continuous pursuit tracking task. *Frontiers in psychology*, 8:2241, 2017. doi: 10.3389/fpsyg.2017.02241 2
- [18] P. Fitts and M. Posner. Human performance. 1967. 2
- [19] 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, 2021. doi: 10.1109/TVCG.2020.2999197 1, 2, 3, 4, 7, 8, 9
- [20] D. T. Gabbett, M. Wake, and B. Abernethy. Use of dual-task methodology for skill assessment and development: Examples from rugby league. *Journal of Sports Sciences*, 29(1):7–18, 2011. PMID: 21086214. doi: 10.1080/02640414.2010.514280 2
- [21] T. J. Gabbett and B. Abernethy. Dual-task assessment of a sporting skill: influence of task complexity and relationship with competitive performances. *Journal of sports sciences*, 30(16):1735–1745, 2012. doi: 10.1080/02640414.2012.713979 2
- [22] S. Gallagher. Body image and body schema: A conceptual clarification. *The Journal of mind and behavior*, pp. 541–554, 1986. 1, 2
- [23] M. Gallotti and C. D. Frith. Social cognition in the we-mode. *Trends in cognitive sciences*, 17(4):160–165, 2013. doi: 10.1016/j.tics.2013.02.002 9
- [24] G. Ganesh, A. Takagi, R. Osu, T. Yoshioka, M. Kawato, and E. Burdet. Two is better than one: Physical interactions improve motor performance in humans. *Scientific reports*, 4(1):1–7, 2014. doi: 10.1038/srep03824 7
- [25] M. Gonzalez-Franco, E. Ofek, Y. Pan, A. Antley, A. Steed, B. Spanlang, A. Maselli, D. Banakou, N. Pelechano, S. Orts-Escolano, et al. The rocketbox library and the utility of freely available rigged avatars. *Frontiers in virtual reality*, 1:561558, 2020. doi: 10.3389/frvir.2020.561558 3
- [26] T. Hagiwara, G. Ganesh, M. Sugimoto, M. Inami, and M. Kitazaki. Individuals prioritize the reach straightness and hand jerk of a shared avatar over their own. *Iscience*, 23:101732, 2020. doi: 10.1016/j.isci.2020.101732 2, 4, 7, 9
- [27] T. Hagiwara, M. Sugimoto, M. Inami, and M. Kitazaki. Shared body by action integration of two persons: Body ownership, sense of agency and task performance. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 954–955. IEEE, 2019. doi: 10.1109/VR.2019.8798222 2
- [28] C. Handford, K. Davids, S. Bennett, and C. Button. Skill acquisition in sport: Some applications of an evolving practice ecology. *Journal of sports sciences*, 15(6):621–640, 1997. doi: 10.1080/026404197367056 1
- [29] H. Hapuarachchi and M. Kitazaki. Knowing the intention behind limb movements of a partner increases embodiment towards the limb of joint avatar. 2022. doi: 10.1038/s41598-022-15932-x 2
- [30] H. Hapuarachchi and M. Kitazaki. Knowing the partner’s objective increases embodiment towards a limb controlled by the partner. In *Proceedings of the 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 802–803, 2022. doi: 10.1109/VRW55353.2022.00252 2
- [31] N. A. Hayes and D. E. Broadbent. Two modes of learning for interactive tasks. *Cognition*, 28(3):249–276, 1988. doi: 10.1016/0010-0277(88)90015-7 2
- [32] J. Higgins. Cognising with others in the we-mode: A defence of ‘first-person plural’ social cognition. *Review of Philosophy and Psychology*, 12(4):803–824, 2021. doi: 10.1007/s13164-020-00509-2 9
- [33] O. Hikosaka, M. Rand, K. Nakamura, S. Miyachi, K. Kitaguchi, K. Sakai, X. Lu, and Y. Shimo. Long-term retention of motor skill in macaque monkeys and humans. *Experimental brain research*, 147(4):494–504, 2002. doi: 10.1007/s00221-002-1258-7 2
- [34] A. Hiyama, Y. Doyama, M. Miyashita, E. Ebuchi, M. Seki, and M. Hirose. Artisan training using wearable egocentric display. In *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology*, pp. 81–82, 2011. doi: 10.1145/2046396.2046433 1
- [35] A. Iriki, M. Tanaka, and Y. Iwamura. Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, 7(14):2325–2330, 1996. doi: 10.1097/00001756-199610020-00010 2
- [36] M. Ito. Mechanisms of motor learning in the cerebellum. *Brain research*, 886(1-2):237–245, 2000. doi: 10.1016/S0006-8993(00)03142-5 7
- [37] C. Jeunet, L. Albert, F. Argelaguet, and A. Lécuyer. “do you feel in control?": towards novel approaches to characterise, manipulate and measure the sense of agency in virtual environments. *IEEE transactions on visualization and computer graphics*, 24(4):1486–1495, 2018. doi: 10.1109/TVCG.2018.2794598 2
- [38] J. Jung and Y. J. Ahn. Effects of interface on procedural skill transfer in virtual training: Lifeboat launching operation study. *Computer Animation and Virtual Worlds*, 29(3-4):e1812, 2018. doi: 10.1002/cav.1812 1
- [39] S. Kager, A. Hussain, A. Cherpin, A. Melendez-Calderon, A. Takagi, S. Endo, E. Burdet, S. Hirche, M. H. Ang, and D. Campolo. The effect of skill level matching in dyadic interaction on learning of a tracing task. In *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, pp. 824–829, 2019. doi: 10.1109/ICORR.2019.8779485 7, 8
- [40] E. Kal, R. Prossé, M. Winters, and J. Van Der Kamp. Does implicit motor learning lead to greater automatization of motor skills compared to explicit motor learning? a systematic review. *PloS one*, 13(9):e0203591, 2018. doi: 10.1371/journal.pone.0203591 2, 3, 8
- [41] M. Kawato, K. Furukawa, and R. Suzuki. A hierarchical neural-network model for control and learning of voluntary movement. *Biological cybernetics*, 57(3):169–185, 1987. doi: 10.1007/BF00364149 7
- [42] J. Keating and W. Thach. Nonclock behavior of inferior olive neurons: interspike interval of purkinje cell complex spike discharge in the awake behaving monkey is random. *Journal of neurophysiology*, 73(4):1329–1340, 1995. doi: 10.1152/jn.1995.73.4.1329 7
- [43] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012. doi: 10.1162/PRES_a_00124 5
- [44] S. T. Klapp, M. D. Hill, J. G. Tyler, Z. E. Martin, R. J. Jagacinski, and M. R. Jones. On marching to two different drummers: perceptual aspects of the difficulties. *Journal of Experimental Psychology: Human Perception and Performance*, 11(6):814, 1985. doi: 10.1037/0096-1523.11.6.814 9
- [45] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi, and M. Hirose. Enhancing the sense of agency by transitional weight control in virtual co-embodiment. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2022. doi: 10.1109/ISMAR55827.2022.

- 00043 8
- [46] T. Kojima, A. Hiyama, T. Miura, and M. Hirose. Training archived physical skill through immersive virtual environment. In *Proceedings of the International Conference on Human Interface and the Management of Information*, pp. 51–58, 2014. doi: 10.1007/978-3-319-07863-2_6 2
- [47] M. L. Latash, J. P. Scholz, and G. Schönner. Motor control strategies revealed in the structure of motor variability. *Exercise and sport sciences reviews*, 30(1):26–31, 2002. 9
- [48] H. M. Lefcourt. *Locus of control*. Academic Press, 1991. 9
- [49] R. Llorens, A. Borrego, P. Palomo, A. Cebolla, E. Noé, S. B. i Badia, and R. Baños. Body schema plasticity after stroke: subjective and neurophysiological correlates of the rubber hand illusion. *Neuropsychologia*, 96:61–69, 2017. doi: 10.1016/j.neuropsychologia.2017.01.007 2
- [50] A. Maravita and A. Iriki. Tools for the body (schema). *Trends in cognitive sciences*, 8(2):79–86, 2004. doi: 10.1016/j.tics.2003.12.008 2
- [51] E. J. A. Míreles, D. De Santis, P. Morasso, and J. Zenzeri. Transferring knowledge during dyadic interaction: The role of the expert in the learning process. In *Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2149–2152, 2016. doi: 10.1109/EMBC.2016.7591154 7
- [52] H. O’shea and A. Moran. Are fast complex movements unimaginable? pupillometric studies of motor imagery in expert piano playing. *Journal of Motor Behavior*, 2018. doi: 10.1080/00222895.2018.1485010 9
- [53] P. N. Parmar, F. C. Huang, and J. L. Patton. Evidence of multiple coordinate representations during generalization of motor learning. *Experimental brain research*, 233(1):1–13, 2015. doi: 10.1007/s00221-014-4034-6 1
- [54] H. Pashler. Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, 116(2):220, 1994. doi: 10.1037/0033-2909.116.2.220 2
- [55] K. Poeck and B. Orgass. The concept of the body schema: A critical review and some experimental results. *Cortex*, 7(3):254–277, 1971. doi: 10.1016/S0010-9452(71)80005-9 1
- [56] C. L. Reed. What is the body schema? 2002. doi: 10.1017/CBO9780511489969.014 2
- [57] D. Roth and M. E. Latoschik. Construction of the virtual embodiment questionnaire (veq). *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3546–3556, 2020. doi: 10.1109/TVCG.2020.3023603 5, 8
- [58] P. Salembier, J. Kahn, G. Calvet, M. Zouinar, and M. Relieu. “just follow me”. examining the use of a multimodal mobile device in natural settings. In *proceedings of the HCI international conference july*, pp. 22–27. Citeseer, 2005. 2
- [59] W. Schneider and A. D. Fisk. Attention theory and mechanisms for skilled performance. In *Advances in psychology*, vol. 12, pp. 119–143. Elsevier, 1983. doi: 10.1016/S0166-4115(08)61989-5 2
- [60] N. Sebanz, H. Bekkering, and G. Knoblich. Joint action: bodies and minds moving together. *Trends in cognitive sciences*, 10(2):70–76, 2006. doi: 10.1016/j.tics.2005.12.009 9
- [61] N. Sebanz, G. Knoblich, and W. Prinz. Representing others’ actions: just like one’s own? *Cognition*, 88(3):B11–B21, 2003. doi: 10.1016/S0010-0277(03)00043-X 9
- [62] T. J. Shuell. Phases of meaningful learning. *Review of educational research*, 60(4):531–547, 1990. doi: 10.3102/00346543060004531 2
- [63] M. Tombu and P. Jolicœur. A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 29(1):3, 2003. doi: 10.1037/0096-1523.29.1.3 2
- [64] W. Wen, A. Yamashita, and H. Asama. The sense of agency during continuous action: performance is more important than action-feedback association. *PLOS ONE*, 10(4):e0125226, 2015. doi: 10.1371/journal.pone.0125226 8
- [65] W. Wen, A. Yamashita, and H. Asama. The influence of performance on action-effect integration in sense of agency. *Consciousness and Cognition*, 53:89–98, 2017. doi: 10.1016/j.concog.2017.06.008 8
- [66] D. B. Willingham. A neuropsychological theory of motor skill learning. *Psychological review*, 105(3):558, 1998. doi: 10.1037/0033-295X.105.3.558 2
- [67] G. Wulf, C. Shea, and R. Lewthwaite. Motor skill learning and performance: a review of influential factors. *Medical education*, 44(1):75–84, 2010. doi: 10.1111/j.1365-2923.2009.03421.x 2
- [68] U. Yang. Just follow me: An immersive vr-based motion training system. In *Proceedings of the International Conference on Virtual Systems and Multimedia*, 1999, pp. 435–444, 1999. 2
- [69] U. Yang and G. J. Kim. Implementation and evaluation of “just follow me”: An immersive, vr-based, motion-training system. *Presence: Teleoperators & Virtual Environments*, 11(3):304–323, 2002. doi: 10.1162/105474602317473240 1, 2
- [70] K. Yun, K. Watanabe, and S. Shimojo. Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific reports*, 2(1):1–8, 2012. doi: 10.1038/srep00959 9