Exploring Perspective Dependency in a Shared Body with Virtual Supernumerary Robotic Arms

1st Ryo Takizawa
Keio University
Kanagawa, Japan
r-takizawa@imlab.ics.keio.ac.jp

2nd Adrien Verhulst
The University of Tokyo
Tokyo, Japan
adrienverhulst@star.rcast.u-tokyo.ac.jp

3rd Katie Seaborn
RIKEN AIP
Tokyo, Japan
katie.seaborn@riken.jp

4th Masaaki Fukuoka
Keio University
Kanagawa, Japan
mskifukuoka@imlab.ics.keio.ac.jp

5th Atsushi Hiyama
The University of Tokyo
Tokyo, Japan
hiyama@star.rcast.u-tokyo.ac.jp

6th Michiteru Kitazaki
Toyohashi University of Technology
Aichi, Japan
mich@tut.jp

7th Masahiko Inami
The University of Tokyo
Tokyo, Japan
inami@star.rcast.u-tokyo.ac.jp

8th Maki Sugimoto
Keio University
Kanagawa, Japan
maki.sugimoto@keio.jp

***Abstract*—With advancements in robotics, systems featuring wearable robotic arms teleoperated by a third party are appearing. An important aspect of these systems is the visual feedback provided to the third party operator. This can be achieved by placing a wearable camera on the robotic arm's "host", but such a setup makes the visual feedback dependent on the movements of the main body. Here we reproduced this view dependency in VR using a shared body. The "host" shares their virtual body with the virtual Supernumerary Robotic Arms of the teleoperator. In our research, the two users perform two tasks: (i) a "synchronization task" to improve their joint action performance and (ii) a "building task" where they worked together to build a tower. In a user study, we evaluated the embodiment, workload, and performance of the teleoperator through the "building task" with three different view dependency modes. We present some early outcomes as trends that might give directions for the future investigations.**

***Index Terms*—Virtual reality, Shared body, Supernumerary Limbs, Third Arm**

I. Introduction

Collaborative Virtual Reality (VR) systems are exploring new forms of interaction between systems and people. One that has gained attention recently is Supernumerary Robotic Limbs (SRLs), commonly referred to as "exoskeletons," with Supernumerary Robotic Arms (SRAs) in particular receiving attention in academia and industry. SRLs/SRAs are a form of "robotic augmentation" [1] useful for reducing physical workload [2] and providing the use of more than two arms. They have been adopted in healthcare [3], and in several other sectors of activities (e.g., construction [4] and fabrication). As an indication of their usefulness, the exoskeletons' market is expected to reach USD $2.5 billion by 2024 [5] (as of 2019).

Cooperation between an **operator** and one or more SRAs can make physical tasks easier but can also be challenging in itself, requiring the operator to balance their arm movements and SRA movements. Several efforts have attempted to resolve this issue, such as by not using limb mapping (e.g. using a joystick instead), remapping the robotic arms to other body parts [6], [7], recording and/or playing commands, or using AI for task recognition [8], [9] Yet all of these strategies have limitations that point towards exploring alternative solutions.

One alternative is to avoid the issue altogether by including a **teleoperator**. Indeed, SRAs can be worn by a host but controlled by one or more teleoperators. For instance, the Fusion system [10] features two SRAs controlled by a teleoperator while a robotic camera "head" attached to the host provides visual feedback. The teleoperator and the host can thus "share a body" and collaborate remotely. However, teleoperation presents challenges that need to be considered and researched from the user's perspective, for instance, the teleoperator lacks of control over its viewpoint and over the shared body.

Here we created and evaluated collaborative SRAs in a shared body. We used a Virtual Environment (VE) with virtual SRAs, since VR is a cost-effective and faster approach compared to real SRAs. We focused on the impact that the teleoperator perspective of the scene had on his/her performance, body ownership, workload, and "team empathy". We detail our system design and report on the results of the evaluation in this paper. The main contributions of this work are the following:

1) The effects of different teleoperator perspectives on performance, embodiment, and workload;
2) The feasibility of using VR to explore teleoperator design problems in robotics (c.f. Fusion [10]);
3) The role of cooperative task training with synchrony; and

978-1-7281-5604-0/19/$31.00 ©2019 IEEE
DOI 10.1109/AIVR46125.2019.00014

4) The usability and user experience of an application of VR body sharing involving two people.

The layout of the paper is as follows: In Sec. II we explore the related work, then in Sec. III we introduce the VE alongside the virtual SRAs and in Sec. IV the experiments details. In Sec. V we present the results, discuss them in Sec. VI and conclude in Sec. VII.

II. Related Works

A. Homonculus Flexibility and Body Ownership

Various studies have been conducted on the remapping of the body. Steptoe *et al.* [11] conducted an experiment using a method of remapping physicality by introducing a tailed body in a virtual environment. This study suggests that sense of ownership of the body increases with synchronization of visual movement. Won *et al.* [12] implemented a third arm that remapped the movements of both hands in a virtual environment. In this experiment, participants adapted to the third arm in less than 10 minutes; they also performed better with an additional arm in VR to complete a task that participants needed to touch three target cubes with three arms, indicating that having multiples limbs can be an advantage. Hoyer *et al.* [13] conducted an experiment by placing a sixth finger between the ring finger and the little finger in a virtual environment. Sense of ownership of the sixth finger was high even if the sixth finger was not moving at all, and became higher when the other fingers were remapped and moved. Feuchtner *et al.* [14] implemented a system that can control remote devices using virtual long arms while maintaining the perception that artificial arms are actually part of the user's own body. They showed that perception of ownership for an elongated virtual arm can be preserved if it looks realistic and is connected to the physical body. What these studies suggest is there is a higher degree of body ownership when there is visuomotor synchrony rather than there is not, but even if there no synchrony the feeling of body ownership is still likely to be high if the additional limb does not move by itself.

B. Self-Location and Perspective

Research has also considered the relationship between viewpoint and embodiment. Self-location, or the experience of where "I" am in space [15], is a key aspect of VR embodiment. Huang *et al.* [16] investigated self-location by comparing body-location and location of the viewpoint. By using a HMD and a stereo camera, the participants watched their own body standing in front of them in the experiment. From the experiment, they stated that not only body-location but also the location of the viewpoint should be taken into consideration in the definition of self-location. Aymerich-Franch *et al.* [17] remapped audio-visual perception and gait to a real robot. Acting as participants, they observed themselves to investigate self-location and self-identification from a first-person epistemological perspective. In their experience, the human mind could exist in two bodies simultaneously.

C. Remote Coordination

Galvan *et al.* [18] proposed how a new method using alternating first person perspective and third person perspective could benefit from specific advantages of each viewpoint. And they said that the interruption of the viewpoint during the simulation suggests that it is not significantly harmful to the sense of body ownership of the virtual body. Nagai *et al.* [19] constructed a system capable of observing and communicating with the entire surrounding visual environment from remote locations by using a wearable camera capable of photographing and transmitting imagery from 360 degrees. They concluded that this may enable simultaneous experiences of travel and professional collaboration with remote places. Komiyama *et al.* [20] constructed a system wherein the user can freely move back-and-forth between first-person and third-person viewpoints in a remote work setting. It enables both the sharing of precise work using first-person images and the understanding of work site conditions using third-person images. Saraiji *et al.* [10] constructed a wearable system that allows operation of two robotic limbs from remote locations for collaborative work and skill sharing. The teleoperator uses a HMD to view the area around the surrogate from the viewpoint of their shoulder. As these studies show, viewpoint is an important factor in remote coordination.

D. Synchronized Action

Synchronized action, where two or more people synchronize their bodily movements in space and time, has been explored in cooperative task settings. In the study by Valdesolo *et al.* [21], two people rocking in chairs at the same time led to increased sensitivity in the perception of the other person. After experiencing the synchronizing pre-task, performance in the cooperative task was improved compared to a non-synchronizing pre-task. Soliman *et al.* [22] investigated how interpersonal synchrony in a pre-task disrupted subsequent independent tasks because participants had adapted to each other, creating an interpersonal joint body schema. These findings suggest that people's tendency to synchronize after a "shared action" experience may lead to synchronization and improved performance in a subsequent joint task carried out in a virtual environment with a virtual shared body. However, this has not yet been explored. Further, there is some evidence to suggest that perception of the other person's psycho-physiology, such as smiling, may be necessary for synchronization to be effective (Mønster *et al.* [23]), but this has not been explored in a VR situation. The context of our study is ideal for exploring the role of interpersonal, especially nonverbal, factors in cooperative situations.

III. System Configuration

The VR system we evaluated allows two participants to share a virtual body, where one is the host, or owner of the main body, and one is a telepresent as an additional limb, and can view the scene from three different perspectives. Here we explain our system design and configuration for this study. Note that from here we will refer to the the host, or **operator**, as the Main Body Operator (Main Body Operator (MBO)) and to the **teleoperator** as the Guest Body Operator (**GBO!** (**GBO!**)).

A. Setup

Participants were tracked in a space that was approximately $10.3m^2$ ($4.37m \times 2.37m$). The MBO skeleton and the **GBO!** arms were resolved with Inverse Kinematics using trackers data (see below) and the FinalIK Unity package[1]. Specifically:

- Tracking of the MBO was performed with three Vive Trackers[2] (waist and legs), two Vive Pro Controllers[3], and a HTC Pro Vive[4].
- Tracking of the Guest Body Operator (GBO) was performed with one Vive Tracker (waist only, there was no need to track the legs), two Vive Pro Controllers, and a HTC Pro Vive.

The application was developed in Unity 3D (v2019) and deployed via two laptops able to run it smoothly. The communication between computers was done with Unity networking via the PUN Unity package[5]; there was no visible latency.

B. Virtual Environment

1) 3D Space: The VE was a semi-realistic rendition of a room that was about the same size as the tracking area (see Sec III-A). It included a TV for displaying short messages (e.g. *"Press the Next button"*) and a "Next button" that both the MBO and the GBO had to select to go to the next step.

2) Main Body and Parasite Body: The GBO was attached to the MBO, so the MBO and the GBO could share the same body. MBO used a dummy model (literally) without a head (so as to not disturb the participant); notably, any avatar, as long as it meets the usual avatar rigging requirements (i.e. having a root and a skeleton rig) can work. The dummy is not a very realistic body, but a realistic body is not necessary to produce a high level of VR embodiment [24].

The GBO was SRAs localized around the MBO's shoulder. The ends of the arms followed the relative position of the GBO's hands, and the remaining part was resolved using IK. Moreover the GBO's default "head position" was about 60cm above the head of the MBO. Its virtual head rotation moves naturally, but the virtual head position moves at a much lower speed (i.e. 20%) than the real head position so as to stay near the MBO.

C. Perspectives Conditions

Three perspectives were experienced by the GBO. These conditions were dependent on the GBO head rotation and position as defined by: (C1) the GBO head rotation and position; (C2) the position of the main body's shoulder AND the GBO head rotation (*only the rotation*); or (C3) the position of the GBO head and the main body's shoulder AND the GBO head rotation. (cf. Tab. I).

[1]http://www.root-motion.com/final-ik.html
[2]https://www.vive.com/fr/vive-tracker/
[3]https://www.vive.com/us/accessory/controller/
[4]https://www.vive.com/us/product/vive-pro/
[5]https://assetstore.unity.com/packages/tools/network/pun-2-free-119922l

TABLE I
PERSPECTIVE CONDITIONS OF THE PARASITE BODY

Depends of	GBO head position	GBO head rotation	Main Body shoulder position
C1	O	O	X
C2	X	O	O
C3	O	O	O

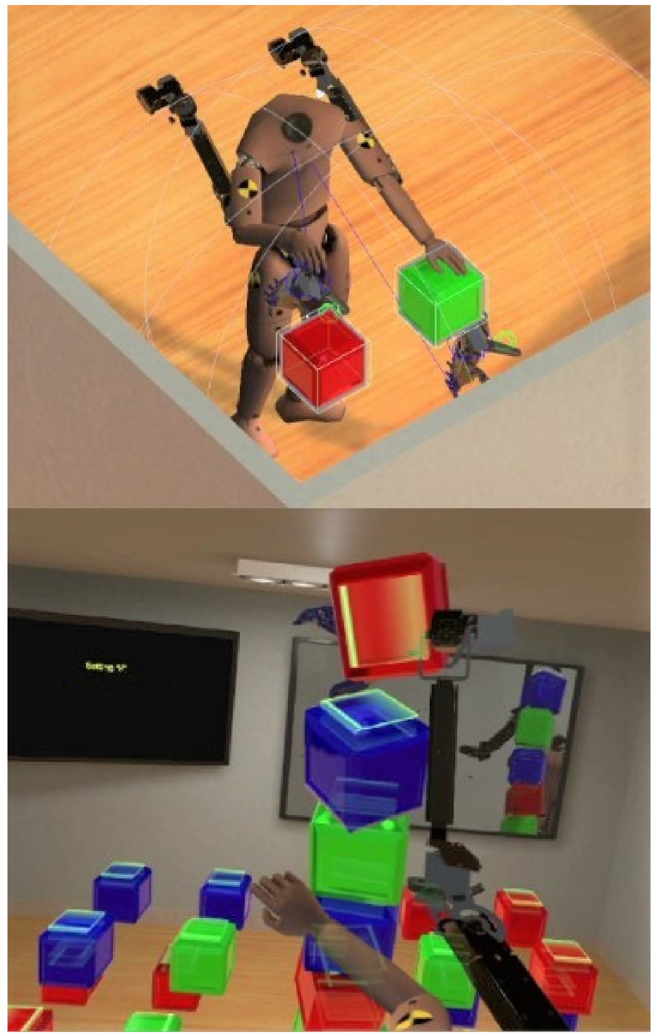

Fig. 1. The experimental tasks. Upper: the Synchronization task. Lower: the Block Tower task.

IV. Methods

We used a mixed experimental design with one between-subjects factor, Synchronization, which had two levels, Synchronized (S) and Non-Synchronized (NS), and one within-subjects factor, Perspective, which had three levels: C1, C2, and C3 cf. III-C. To avoid order effects, the order of C1, C2, and C3 was counterbalanced between participants.

A. Hypotheses

1) Perspective: Depending on the type of task (e.g. if the MBO moves a lot), the motion of the MBO can make it difficult for the GBO to have a good understanding of their location and movements.

> **[H1]** The perspective will affect the GBO's body ownership, workload, and performance.
>
> **[H2]** The perspective will affect the usability of the SRA.

2) Shared Body: When using a shared body, users do not have full control over all of their limbs, thus partially breaking aspects of visuomotor synchrony. Still, different body can be mastered relatively fast and it may be possible to improve performance offsets or performance in general in a shared body context through training that involves synchrony, in line with previous research [21].

> **[H3]** The GBO will be able to quickly adapt to and have good body ownership over the SRA.
>
> **[H4]** A training task based on synchrony will increase shared body performance between the GBO and MBO.

B. Participants

Fifteeen participants acted as the GBO (14 men; age: M = 24.9, SD = 3.0). All were students or faculty members, among which $\approx$ 20% had some previous experience with VR (3-10 times), and $\approx$ 80% had extensive experience with VR (>10 times).

C. Procedure

Participants were invited one-by-one to join the experiment as the GBO and the MBO (after finishing the experiment, they changed the role). Participants read a description of the controls and tasks, then were fitted with the body trackers and HMD and led to the tracking area (cf. Fig 2). Throughout the experiment, the assistant provided help if needed. Participants were asked if they wanted to take a break or stop if there was any sign of discomfort (such as dizziness).

Once in the Virtual Environment (VE), participants completed the "Synchronization" task and then filled out the synchronization questionnaire (step I, cf. Sec. IV-C1). Next, they completed the Block Tower task and filled out a post-task questionnaire (step II, cf. Sec. IV-C2). They repeated this step for each condition: C1, C2, and C3. Once all conditions were completed, they answered a post-test questionnaire (step III, cf. Sec. IV-C3).

1) Step I - The Synchronization task: The participants appear next to two blocks (one for the MBO, one for the GBO), in front of a mirror (so that they can see each other at any time), cf. Fig. 1. Participants had to touch their respective block in rhythm with beats (heard through their HMD headphones). Each time a block was touched, it would change color. In the *Synchronized* condition, the tempo of beats was the same for both, while in the *Non-Synchronized* condition, the tempo of the beats was different for each.

Fig. 2. Participants in the VE. The participant in the back participant was the MBO, and the participant in the front was the GBO.

The task lasted 120s. On completion, the participants removed their HMD and filled out the synchronization questionnaire on a computer (cf. Tab. II. The synchronization questionnaire follows Valdesolo *et al.* [21], particularly its similarity and liking items, and included distractor items [here, S1, S2, and S4]).

TABLE II
THE SYNCHRONIZATION (S) QUESTIONNAIRE

ID	Item
S1	How would you rank your ability to keep rhythm in this task?
S2	How much did you enjoy the ambiance of the room?
S3	To what extent do you feel similar in personality to the other person?
S4	How much do you like rhythm tasks, like this one?
S5	How much do you like the other person?

2) Step II - The Block Tower task: The participants placed blocks one on top of the other to assemble the "longest" tower possible (the tower doesn't have to be vertical), cf. Fig. 1. There were three types of blocks (red blocks, green blocks, and blue blocks), but: (1) each block could only be placed onto a block of a different color (e.g., a red block on top of a red block was not allowed); (2) only the MBO could manipulate the blue blocks, and only the GBO could manipulate the red blocks, but both could manipulate the green blocks. The presence of blocks that cannot be manipulated by each operator is intended to induce collaboration.

The task lasted 160s. On completion, the participants removed their HMD and filled out a post-task questionnaire made up of the following instruments and items:

- A body ownership questionnaire inspired by Gonzalez-Franco *et al.* [25] (using a 7-point Likert response scale), cf. Tab. III;

- A team effectiveness questionnaire, cf. Tab. III, from which the items T1, T2, and T3 were taken fom [26];
- The System Usability Scale (SUS) questionnaire, focused on the view mode (using a 7-point Likert response scale);
- The NASA TLX [27] in its raw form (a.k.a., the Raw TLX, or RTLX).

TABLE III
THE BODY OWNERSHIP (BO) AND TEAM (T) QUESTIONNAIRE

ID	Item
BO1	I felt as if the virtual body was my body
BO2	It felt as if the virtual body I saw was someone else
BO3	It seemed as if I might have more than one body
BO4	It felt like I could control the virtual body as if it was my own body
BO5	The movements of the virtual body were caused by my movements
BO6	I felt as if the movements of the virtual body were influencing my own movements
BO7	I felt as if the virtual body was moving by itself
BO8	I felt as if my body was located where I saw the virtual body
BO9	I felt out of my body
T1	Our team has a meaningful, shared purpose
T2	We are able to resolve conflicts with each other
T3	We addressed and resolve issues quickly
T4	I felt like the other person was disturbing me
T5	I felt like the other person was not helping me

3) Step III - Post-Test Questionnaire: After completing all tasks and all conditions, participants filled out post-test questionnaire comprised of the following items:

- Basic demographics, including sex, age, ethnicity, and experience with VR;
- The Presence Questionnaire from Usoh *et al.* [28] (using a 7-point Likert response scale);
- Four open questions about the experiment (*"Please write at least 2 sentences about what you felt or thought when you were controlling the avatar."*, *"Were the questions clear enough?"*, *"Please write at least 2 sentences about what you like / did not like in the experiment, as well as what we could improve in the experiment."*, *"Please write what you think the purpose of the experiment was"*).

D. Performance Metrics

We also recorded the number of blocks touched as well as the number of blocks placed by the GBO.

V. RESULTS

Here we present the results of our study.

A. Body Ownership

Since the body ownership questionnaire items were heavily adapted for our circumstances, we first verified their internal consistency by computing Cronbach's alpha coefficients. The result for the 8 items was 0.72 (internally consistent).

We calculated the body ownership score for the MBO (overall $M = 49.4, SD = 6.16$) and the GBO (cf. Tab. IV). We also computed Friedman test for the GBO to see if there were any significant differences. There was a significant difference for the item BO4 and BO6. Comparisons of paired t-tests for BO4 showed no significant difference between conditions, and for BO6 showed a significant difference between C1 and C2 ($t = -2.39, df = 15, p = 0.03$), with means indicating that C1 was significantly worse than C2.

TABLE IV
BODY OWNERSHIP VALUES FOR THE GBO. NEGATIVE ITEMS WERE REVERSED.

Items	C1		C2		C3		Friedman	
	M	SD	M	SD	M	SD	χ^2	p
BO1	4.12	1.54	4.31	1.82	4.20	1.57	1.03	0.60
BO2	4.25	1.34	4.25	1.81	3.27	1.49	4.44	0.11
BO3	4.31	1.40	4.94	1.39	3.40	1.35	3.77	0.15
BO4	3.81	1.64	3.56	1.82	4.00	1.56	6.9	**0.03**
BO5	5.19	1.17	4.81	1.76	5.27	1.03	0.70	0.70
BO6	2.56	1.03	3.50	1.59	2.67	1.05	6.28	**0.04**
BO7	4.31	1.40	4.00	1.59	3.93	1.62	0.32	0.85
BO8	3.50	1.83	4.56	1.41	3.60	1.55	0.17	0.91
BO9	2.88	1.67	3.94	1.88	2.87	1.51	1.58	0.45
BO Score	34.94	7.37	37.88	8.74	33.20	7.24	1.85	0.40

B. RTLX

We calculated the RTLX score for the MBO (overall $M = 8.23, SD = 4.12$) and the GBO (cf. Tab. V). We also computed a Friedman test for each item of the RTLX for the GBO, but there were no significant differences .

TABLE V
THE RTLX VALUES FOR THE GBO

Items	C1		C2		C3		Friedman	
	M	SD	M	SD	M	SD	χ^2	p
Mental	11.81	5.06	11.88	5.25	11.93	4.92	0.37	0.83
Physical	11.25	5.27	11.75	4.84	11.00	4.81	0.86	0.65
Temporal	5.56	3.86	5.06	3.82	5.47	3.98	0.06	0.97
Performance	10.69	5.61	9.50	4.79	10.73	4.99	0.76	0.68
Effort	13.00	5.06	12.50	4.72	13.07	4.18	0.90	0.64
Frustration	9.31	5.91	10.31	5.35	10.47	4.24	2.40	0.30
RTLX	10.27	3.95	10.17	3.55	10.44	3.28	0.00	1.00

C. Team Effectiveness and System Usability

We calculated a team empathy score for the GBO (cf. Tab. VI). We also computed a Friedman test for each item, but there were no significant differences (with $F(2, 28)$).

We also calculated the SUS score for the PBO and then computed a Friedman, but there was no significant differences (with $F(2, 28)$).

TABLE VI
THE TEAM VALUES AND SUS FOR THE GBO

Items	C1		C2		C3		Friedman	
	M	SD	M	SD	M	SD	χ^2	p
T1	5.75	1.06	6.00	0.82	5.67	1.18	0.06	0.97
T2	5.25	0.93	5.38	1.45	5.20	1.26	0.20	0.90
T3	5.44	1.82	5.75	1.69	6.20	1.01	2.21	0.33
T4	5.50	1.51	5.12	1.45	5.80	1.15	6.50	0.06
T5	6.00	0.97	5.56	1.41	6.07	0.80	3.52	0.17
SUS	21.94	6.72	22.75	5.12	22.0	6.63	0.33	0.85

D. Performance

We calculated performance values for the GBO (cf. Tab. VII). We also computed a ANOVA for each item (with $F(2, 28)$). We found a significant difference for the number of red blocks touched, the total of blocks placed, and the green blocks missed. Paired t-tests comparing these items found the following: (i) for red blocks touched, significant differences between C1 and C3 ($t = 3.24, df = 14, p = 0.006$) and C2 and C3 ($t = -2.17, df = 14, p = 0.05$), with C3 better than C1 and C2; (ii) for total blocks placed, no significant differences; and (iii) for green blocks missed, a significant difference between C2 and C3 ($t = 2.98, df = 14, p = 0.01$), with C3 better than C2.

TABLE VII
THE PERFORMANCES VALUES FOR THE GBO

Items	C1		C2		C3		ANOVA	
	M	SD	M	SD	M	SD	F	p
Red touch	4.38	2.53	4.62	3.40	6.13	3.16	5.20	**0.01**
Green touch	4.12	2.53	5.25	2.98	4.07	2.49	2.06	0.15
Total touch	8.50	4.26	9.88	5.88	10.20	4.31	2.32	0.12
Red place	2.12	1.54	1.94	1.53	2.53	1.73	1.45	0.25
Green place	1.56	1.50	1.69	1.30	2.40	1.80	2.29	0.12
Total place	3.69	2.75	3.62	2.39	4.93	2.37	4.83	**0.02**
Red miss	2.25	1.65	2.69	2.41	3.60	3.00	1.99	0.16
Green miss	2.56	1.67	3.56	2.34	1.67	1.80	3.59	**0.04**
Total miss	4.81	2.34	6.25	4.25	5.27	3.35	1.40	0.26

E. Presence

We calculated the Presence for the MBO and the GBO (cf. Tab. VIII). We also computed a Kruskal Wallis test for each item, but there were no significant differences.

F. Open Questions

We highlight common patterns of comments reported by participants: • "It's difficult to control the robot arm." (mentioned 9 times) • "I felt ownership of the robot arm." (mentioned 3 times) • "When the viewpoint is dependent on the main body, I feel motion sickness." (mentioned 4 times); • I felt that the cooperation task was clear and interesting. (mentioned 4 times). In general, most participants felt it was difficult to control the robot arm, but some felt ownership of the arm. Regardless, these comments indicate that people can get used to the arm over time.

TABLE VIII
THE PRESENCE QUESTIONNAIRE FOR THE MBO AND THE GBO

Items	Main		Parasite		Kruskal Wallis	
	M	SD	M	SD	χ^2	p
P1	5.89	0.93	5.33	1.80	0.24	0.62
P2	4.00	1.58	3.87	1.46	0.06	0.81
P3	6.11	0.93	5.73	1.33	0.28	0.59
P4	3.33	2.18	2.47	1.41	0.67	0.42
P5	3.78	1.72	4.33	1.29	0.68	0.41
P6	4.78	1.39	4.27	1.79	0.37	0.54
P Score	27.67	6.50	26.07	5.76	0.29	0.59

VI. DISCUSSION

A. Regarding H1 and H2

We wanted to know if perspective would affect body ownership, workload, and performance (H1, cf. Sec. IV-A). However, we did not find any significant differences in the "Body ownership" scores (except BO6, discussed below), nor in the RTLX scores. Still, we found significant differences in performance scores, to be discussed.

There was no difference in usability between conditions, meaning H2 is not verified. However let us add that the SUS score were extremely low (above 61.2 is a "above average" usability and below 61.2 is a "below average" usability). Our system SRA did not follow the arm movement like a "skeleton" do, but was following a IK target, therefore it was indeed relatively difficult to use.

B. Body Ownership

Regarding the body ownership score, we could find a significant difference only in the BO6 "I felt as if the movements of the virtual body were influencing my own movements". Since the viewpoint of C2 depends on the shared body compared to C1, it seems that the ownership of this body has increased. If there is more time to get used to, we can find more significant differences.

C. Workload

Regarding the RTLX score, we did not find any significant differences. We realize that this is in contradiction with our explanation in Sec. VI-G, where participants in C1 had to realign themselves, which should have led to greater workload scores (on the frustration or mental scales, for example). In fact, our results are nearly identical in each condition. A possible explanation is that the participants felt some level of difficulty in every task, and did not have time to get used to it (as they were focused on the task) to appreciate the differences. Retrospectively, a training session involving the different viewpoints would have been beneficial. Another explanation (although without relation to the results) is that we used the RTLX instead of the *weighted* NASA TLX; perhaps with weights, the scores would have been different.

Overall, the workload was average, with the "Effort" item being the highest in every condition (the task required some effort, as it was necessary to do multiple actions at the same time, including move around, be attentive to the MBO, etc.). The "Temporal" item was the lowest; while the task was limited in time, participants were not aware of the elapsed time, so there was no real incentive to go as fast as possible. Rather, participants took the time to help each other and make sure they each had the opportunity to place their own blocks).

D. Performance

In terms of performance, C3 allowed for the touching of significantly more red blocks than in C1 and C2. Note that we changed the order of the conditions, therefore it is not because in C3 the participants had accumulated more experience with the task. Although the average of blocks touched was also higher for C3, it was not significantly so. We do not have an explanation as to why the participants touched more red blocks than green blocks in C3. Similarly, participants in C2 missed significantly more green blocks than in C3. Again, there is no definitive explanation. However, taken together, and also considering the fact that the average participant in C3 placed more blocks than in C1 and C2, hint that C3 allowed for better overall performance, with a higher chance of grabbing a block and a lower chance of missing its placement.

E. Team Effectiveness

In the pilot study, some GBO participants reported that the MBO was "disturbing" them. To explore this further, we used a Team questionnaire to know more about the feeling of the GBO. The results did not provide any significant differences and indeed showed an overwhelming appreciation of the other participant. However, we would ask that the reader to not read too much into these high numbers as: (i) the MBO and GBO knew each other beforehand in every case, and (2) the participants were seated next to each other when answering the questionnaires (they could not see each other answer, but they were within arm's distance).

F. Regarding H3: Embodiment Adaptation

Overall, the "Embodiment" score was above average ($average = 31.5$) in all cases. For the MBOs, it was far above average ($M = 49.4$), showing that our setup was sufficient to have body ownership. The GBOs' scores were also in all conditions above average.

G. Embodiment

Regarding BO6, *"I felt as if the movements of the virtual body were influencing my own movements"*, C1 was significantly worse than C2. In C2, the GBO cannot move their virtual body at all (the movements are dependant on the MBO). We believe that in C1, when the MBO moves, they force the GBO to re-position themself. We noticed that it is difficult (cognitively) to manipulate the SRA if we are not oriented like the SRA (i.e., facing the same forward direction). It is therefore not surprising to see that in C1 the GBO is influenced by the movements of their virtual body (i.e., the virtual body movements >my viewpoint does not follow the movement, therefore I am not facing the same direction as the SRA >I move myself to be in the same direction as the SRA).

The lack of significance in the "Embodiment" score is harder to explain. We assumed that not being able to move your virtual body (C2) versus being able to (C1) or even partially able to (C3) would have had a greater impact. However, EM3 *"It seemed as if I might have more than one body"* and EM9 *"I felt out of my body"* were close to significant, so it is possible that with more participants we would have been able to see more clearly defined effects.

H. Presence

Presence was above average (> 21) for both the GBO and the MBO, with no significant difference among items. This suggests that the teleoperation context did not "break" or reduce the feeling of presence in VR.

I. Regarding H4: Training with Synchrony

No significant differences were found between the synchrony and non-synchrony groups. This goes against expectations based on the previous research [21], [22], where a pre-task based on synchronous action resulted in higher cooperative performance in a subsequent main task. Synchrony may require visual perception of the synchronous actions; since the point of view differed between the perspective conditions, which all participants experiences, the effect may have been lost, especially as participants became used to task (which repeated). This may also provide initial evidence for the value, if not necessity, of psycho-physiological information about the other person, which is not available in this VR context [23].

J. Limitations

The statistical analyses must be interpreted with care, given the experimental design and low sample size. This study is also limited by the inclusion of young people; it is not known whether and how the experience of the three perspective conditions changes for other age groups. Similarly, most of the participants are male, so it is very unbalanced. Also, the collision between MB arms and PB arms is not considered, but is important when considering real-world context.

VII. Conclusion

In this work, we investigated the effects of different perspectives of a teleoperator on performance, VR embodiment and body ownership, workload, and usability, and explored an application of VR body sharing. We cannot provide a definitive answer for H1 (cf. Sec. IV-A). However, we showed that there was a significant difference in the embodiment, that the workload was essentially the same for this task, and that some aspects of performance varied based on perspective, hinting towards condition C3 being superior in this context. H2 about the usability (cf. Sec. IV-A), was verified since there were significant differences in the performance. H3, or ease and speed of body adaptation (cf. Sec. IV-A), was verified. We also found no difference in presence between the GBO and the MBO and in the "team" items. Finally, it is unclear what effect, if any, synchrony had on improving cooperation in a VR shared body, and more research is needed focusing on this aspect. We believe that this work is a step forward in research on shared body contexts in VR and provides some insight for non-virtual, "real-life" applications like teleoperated SRAs. Overall, these results hint that the perspective dependency has little incidence regarding Body Ownership and Workload, and some regarding the task's performance. We would say that viewing through a camera dependant of the movements of another person still allow the teleoperator to feel Body Ownership and to complete shared actions.

REFERENCES

[1] Sang-won Leigh, Harshit Agrawal, and Pattie Maes. Robotic Symbionts: Interweaving Human and Machine Actions. *IEEE Pervasive Computing*, 17(2):34–43, apr 2018.

[2] David Allison. Delta air lines may outfit some employees with 'wearable robotics'. http://www.root-motion.com/final-ik.html, Mar 2018. (accessed 2019-7-10).

[3] Science Service Dr. Hempel GmbH. Helping people walk with exoskeletons — the growth of wearable robots in healthcare. https://www.dr-hempel-network.com/digital-health-technolgy/wearable-robots-in-healthcare/, Jul 2018. (accessed 2019-7-10).

[4] Federico Parietti and Harry Asada. Supernumerary Robotic Limbs for Human Body Support. *IEEE Transactions on Robotics*, 32(2):301–311, 2016.

[5] John Bay. Exoskeleton market to surpass us$2.5 billion by 2024. https://www.marketresearchengine.com/exoskeleton-market/, May 2019. (accessed 2019-7-10).

[6] Elahe Abdi, Etienne Burdet, Mohamed Bouri, Sharifa Himidan, and Hannes Bleuler. In a demanding task, three-handed manipulation is preferred to two-handed manipulation. *Scientific Reports*, 6:1–11, 2016.

[7] Adam Drogemuller, Adrien Verhulst, Benjamin Volmer, Bruce Thomas, Masahiko Inami, and Maki Sugimoto. Remapping a third arm in virtual reality. In *Osaka, IEEE VR 2019: the 26th IEEE Conference on Virtual Reality and 3D User Interfaces*, 02 2019.

[8] Baldin Llorens Bonilla and H Harry Asada. A robot on the shoulder: Coordinated human-wearable robot control using Coloured Petri Nets and Partial Least Squares predictions. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 119–125. IEEE, may 2014.

[9] Benjamin Volmer, Adam Drogemuller, Adrien Verhulst, Benjamin Volmer, Bruce Thomas, Masahiko Inami, and Maki Sugimoto. Towards robot arm training in virtual reality using partial least squares regression. In *Osaka, IEEE VR 2019: the 26th IEEE Conference on Virtual Reality and 3D User Interfaces*, 03 2019.

[10] MHD Yamen Saraiji, Tomoya Sasaki, Reo Matsumura, Kouta Minamizawa, and Masahiko Inami. Fusion: Full body surrogacy for collaborative communication. In *ACM SIGGRAPH 2018 Emerging Technologies*, SIGGRAPH '18, pages 7:1–7:2, New York, NY, USA, 2018. ACM.

[11] W. Steptoe, A. Steed, and M. Slater. Human tails: Ownership and control of extended humanoid avatars. *IEEE Transactions on Visualization & Computer Graphics*, 19(4):583–590, April 2013.

[12] Andrea Stevenson Won, Jeremy Bailenson, Jimmy Lee, and Jaron Lanier. Homuncular flexibility in virtual reality. *Journal of Computer-Mediated Communication*, 20(3):241–259, 2015.

[13] Ludovic Hoyet, Ferran Argelaguet, Corentin Nicole, and Anatole Lécuyer. "wow! i have six fingers!": Would you accept structural changes of your hand in vr? *Frontiers in Robotics and AI*, 3(27), 2016.

[14] Tiare Feuchtner and Jörg Müller. Extending the body for interaction with reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5145–5157, New York, NY, USA, 2017. ACM.

[15] Niclas Braun, Stefan Debener, Nadine Spychala, Edith Bongartz, Peter Sörös, Helge H.O. Müller, and Alexandra Philipsen. The senses of agency and ownership: A review. *Frontiers in Psychology*, 9(April):1–17, 2018.

[16] Hsu-Chia Huang, Yen-Tung Lee, Wen-Yeo Chen, and Caleb Liang. The sense of 1pp-location contributes to shaping the perceived self-location together with the sense of body-location. *Frontiers in Psychology*, 8:370, 2017.

[17] Laura Aymerich-Franch, Damien Petit, Gowrishankar Ganesh, and Abderrahmane Kheddar. The second me: Seeing the real body during humanoid robot embodiment produces an illusion of bi-location. *Consciousness and Cognition*, 46:99 – 109, 2016.

[18] Henrique Galvan Debarba, Sidney Bovet, Roy Salomon, Olaf Blanke, Bruno Herbelin, and Ronan Boulic. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PLOS ONE*, 12(12):1–19, 12 2017.

[19] Shohei Nagai, Shunichi Kasahara, and Jun Rekimoto. Livesphere: Sharing the surrounding visual environment for immersive experience in remote collaboration. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '15, pages 113–116, New York, NY, USA, 2015. ACM.

[20] Ryohei Komiyama, Takashi Miyaki, and Jun Rekimoto. Jackin space: Designing a seamless transition between first and third person view for effective telepresence collaborations. In *Proceedings of the 8th Augmented Human International Conference*, AH '17, pages 14:1–14:9, New York, NY, USA, 2017. ACM.

[21] Piercarlo Valdesolo, Jennifer Ouyang, and David DeSteno. The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology*, 46(4):693 – 695, 2010.

[22] Ferguson R. Dexheimer M. S. & Glenberg A. M. Soliman, T. M. Consequences of joint action: Entanglement with your partner. *Journal of Experimental Psychology: General*, 144:873–888, 2015.

[23] Dan Mønster, Dorthe Døjbak Håkonsson, Jacob Kjær Eskildsen, and Sebastian Wallot. Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiology & Behavior*, 156:24 – 34, 2016.

[24] Jean-Luc Lugrin, Maximilian Landeck, and Marc Erich Latoschik. Avatar embodiment realism and virtual fitness training. In *2015 IEEE Virtual Reality (VR)*, pages 225–226. IEEE, mar 2015.

[25] Mar Gonzalez-Franco and Tabitha C. Peck. Avatar Embodiment. Towards a Standardized Questionnaire. *Frontiers in Robotics and AI*, 5(June):1–9, jun 2018.

[26] Team effectiveness questionnaire. https://www.londonleadershipacademy.nhs.uk/leadershiptoolkit/find. (accessed 2019-7-10).

[27] Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in Psychology*, pages 139–183. Elsevier, 1988.

[28] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments*, 9(5):497–503, October 2000.