The Supernumerary Hand Illusion in Augmented Reality

NINA ROSA, REMCO C. VELTKAMP, and WOLFGANG HÜRST, Utrecht University, The Netherlands
TANJA NIJBOER, Utrecht University and University Medical Center Utrecht and Rehabilitation Center De Hoogstraat, The Netherlands
CAROLIEN GILBERS, Utrecht University, The Netherlands
PETER WERKHOVEN, Utrecht University and TNO

The classic rubber hand illusion (RHI) experiment studies the sense of embodiment over a fake limb. Distinguished subcomponents of embodiment are ownership (sense of self-attribution of a body), agency (sense of having motor control), and self-location (the spatial experience of being inside a body), and are typically evoked in either reality or virtual reality. In augmented reality (AR), however, visually present real limbs can be augmented with (multiple) fake virtual limbs, which results in a variation of the RHI, the augmented reality supernumerary hand illusion (ARSHI). Such conditions occur, for example, in first-person AR games and in AR-interfaces for tele-robotics. In this article, we examined to what extent humans can experience the sense of embodiment over a supernumerary virtual arm in addition to one or two real arms. We also examine how embodiment is affected by the perceptual visual-tactile synchronicity of the virtual and real limbs, and by the synchronicity of active movement of the virtual and real hand. Embodiment was measured subjectively by questionnaire and objectively by skin conductance responses (SCRs). Questionnaire responses show that ownership, agency, and self-location can be evoked over the virtual arm in the presence of a real arm, and that they are significantly stronger for synchronous conditions than for asynchronous conditions. The perceptual and motorical synchronous condition with three visible hands led to an experience of owning the virtual hand. These responses further show that agency was also strongly experienced over the supernumerary virtual arm, and responses regarding self-location suggest a shift in sensed location when one real arm was in view and an additional location when both real arms where in view. SCRs show no significant effect of condition, but do show a significant habituation effect as a function of the number of conditions performed by participants. When analyzing the relations at the individual participant level between the questionnaire data and skin conductance, we found two clusters of participants: (1) participants with low questionnaire responses and low-medium SCRs and (2) participants with high questionnaire responses and low-high SCRs. Finally, we discuss how virtual hand appearance/realism and willingness to accept virtual limbs could play an important role in the ARSHI, and provide insights on intricacies involved with measuring and evaluating RHIs.

CCS Concepts: • **Human-centered computing → Mixed/augmented reality**; **Empirical studies in HCI**;

This work was supported by the Netherlands Organization for Scientific Research (NWO), project number 022.005.017.
Authors' addresses: N. Rosa, Utrecht University, Department of Information and Computing Sciences, Princetonplein 5, Utrecht, 3584CC, The Netherlands; email: n.e.rosa@uu.nl; R. C. Veltkamp, Utrecht University, Department of Information and Computing Sciences, The Netherlands; email: r.c.veltkamp@uu.nl; W. Hürst, Utrecht University, Department of Information and Computing Sciences, The Netherlands; email: huerst@uu.nl; T. Nijboer, Utrecht University, Department of Psychology, University Medical Center Utrecht, Rehabilitation Medicine, Rehabilitation Center De Hoogstraat, The Netherlands; email: t.c.w.nijboer@uu.nl; C. Gilbers, Department of Information and Computing Sciences, Utrecht University, The Netherlands; P. Werkhoven, Utrecht University, Department of Information and Computing Sciences, TNO, Technical Sciences, The Netherlands; email: p.j.werkhoven@uu.nl.

1544-3558/2019/08-ART12 $15.00
https://doi.org/10.1145/3341225

Additional Key Words and Phrases: Sense of embodiment, rubber hand illusion, virtual hand illusion, supernumerary hand illusion, multisensory, ownership, agency, self-location, augmented reality, skin conductance habituation

ACM Reference format:
Nina Rosa, Remco C. Veltkamp, Wolfgang Hürst, Tanja Nijboer, Carolien Gilbers, and Peter Werkhoven. 2019. The Supernumerary Hand Illusion in Augmented Reality. *ACM Trans. Appl. Percept.* 16, 2, Article 12 (August 2019), 20 pages.
https://doi.org/10.1145/3341225

1 INTRODUCTION

Research over the past 20 years has shown that the experience of what constitutes 'my body' is very malleable. The classic rubber hand illusion (RHI) experiment by Botvinick and Cohen clearly demonstrated the susceptibility to accept fake limbs [3], and many variations of this experiment have appeared since. In their experiment, participants hide their left arm behind a screen and view a rubber hand/arm model placed in front of them. The experimenter strokes the real and rubber hands with a paintbrush for 10 minutes, in one case synchronously, and in another case asynchronously. Through questionnaires and proprioceptive drift measurement (i.e., the feeling of displacement of the real hand location), the authors report that the participants that experienced synchronous stroking felt as if the rubber hand was their own, and that this did not occur in the asynchronous case. In other words, it was possible to evoke a sense of ownership over an "alien limb". It is concluded that there is a three-way interaction between vision, touch, and proprioception, and that intermodal matching can be sufficient for self-attribution.

Ownership, which was studied in this classic experiment, is just one of several important aspects of a concept called *embodiment*. Kilteni et al. [14] define the sense of embodiment towards a body as "the sense that emerges when [the body's] properties are processed as if they were the properties of one's own biological body". Typically, the sense of embodiment is associated with three main concepts [14, 18]:

Ownership. The self-attribution of a body;
Self-Location. The spatial experience of being inside a body; and
Agency. The sense of having motor control.

Various RHI studies have shown that, although these are separate concepts, they are to a high extent influenced by one another and/or by identical contributing factors, such as visual perspective and visuomotor correlation. Key here is multisensory synchrony: in daily life, humans receive multiple signals over various senses that together provide information on what is happening in and outside of their body. This processing quality can be exploited, in which case the mind can be convinced that the body has changed somehow; in the RHI, the mind is convinced that the parts that belong to the body have changed.

The focus of this article concerns two developments in RHI research. First, although the classic RHI experiment was first examined in reality, more and more studies have shown its effectiveness with virtual arms and full bodies in virtual reality (VR), i.e., completely virtual environments. On the other hand, reports of such illusions are sparse in augmented reality (AR), despite the recent popularity of both VR and AR. AR can be defined as a real environment in which we integrate virtual elements, and where the integration is not limited to the visual sense. In an RHI setting, this means there would no longer be a rubber hand, but a virtual hand while the real body remains visible; any virtual limbs would thus be added to the existing limbs. This brings us to the second development, namely the supernumerary hand illusion (SHI), a novel variation of the RHI. In the SHI, either the real hand remains in sight while a fake hand is also present, or the real hand is out of sight while two fake hands of the same laterality are visible. The findings of this illusion in reality and VR are quite mixed, and such illusions in AR are not as common. Together, these developments result in two major issues. First, it is still unclear whether a sense of embodiment can occur over more than our real limbs, i.e., the feasibility of the SHI is still questionable. Second, it is uncertain whether the type of media, such as reality versus AR, for example, changes the extent of the illusion, or whether adaptions of previously known RHI factors are required to achieve this. The goal of this article is to study these issues. We first describe previous works on the SHI in reality, then previous works on the RHI in AR.

1.1 Supernumerary Hand Illusions in Reality

In the following, we first discuss the mixed SHI results to date, and then a number of possible reasons for the seemingly contrasting literature. Armel and Ramachandran [2] implicitly study the SHI when examining the classic RHI and a so-called 'table-illusion' where participants informally reported feeling ownership over a tabletop. An added contribution of their work was the use of skin conductance response (SCR) after a threat to the hand as an objective measure of the illusion. Using the visual presence of the real hand as a control condition, the authors reported significant differences in both subjective illusion rating and objective SCR that together indicate that the illusion can occur when the real hand is hidden, that it cannot occur when the real hand is visible, and that it in fact does not objectively occur for a table.

Other studies have found that participants can still feel ownership over a fake limb when the real hand is visibly present. Schaefer et al. [25] asked participants to look at both of their real hands and a rubber left hand, and stimulated the pinky and thumb of the real left hand by synchronous touching; all other hands were unstimulated, even the fake hand. The authors measured subjective experience, and the angular distance between activated areas in the somatosensory cortex corresponding to the pinky and thumb. They found that when the rubber hand seemed connected to the body, the angular difference was smaller than when they seemed unconnected and when there was no rubber hand at all. These results are supported by similar significant differences in the illusion ratings; however, it must be noted that the overall ratings were quite low. This shift in cortical representation of the thumb indicates that the somatosensory homunculus reflects the perceived shape of the body rather than the physical aspects of peripheral stimulation as previously thought, and in turn that it is in fact the shape of the body that changes. Another study by Guterstam et al. [11] demonstrated that ownership can indeed be induced over a rubber right hand when the real right hand is visible without disownership of the real hand, but also less ownership of the rubber hand compared to a RHI with hidden real right hand. By means of questionnaires and SCR, they showed that by using synchronous stroking a higher level of ownership was achieved over a rubber arm compared to asynchronous stroking, and that ownership did not occur for other limbs or arms of incorrect laterality. Finally, Chen et al. [4] showed that when participants sat across from the experimenter and saw four hands, with the experimenter's hands in first person perspective and their own in third, which they tapped on the table surface synchronously while being synchronously stroked, they subjectively had the experience of owning four hands. This illusion was not reflected in the objective SCRs, and did not occur for only brushing or only tapping, or when the participant saw their own hands in first person perspective.

A last category of SHI studies have used two fake hands while the real hand is hidden. Ehrsson [6] used two rubber hands that were placed above and slightly to the left/right of the real right hand, with the left hand in view. The results showed that there was a significant difference in SCR between synchronous and asynchronous stroking, revealing that healthy individuals are capable of feeling and seeing supernumerary limbs. Newport et al. [21] explore how the SHI affects both the body schema and body image. The authors use two video displays of the real left hand, which is in turn out of sight, and synchronously stroke either the rightmost or leftmost hand (other asynchronously), or both synchronously with the real hand. The effects on the body image were measured by means of a questionnaire, and the effects on the body schema using a pointing task. The questionnaire showed that participants experienced ownership over the hand/hands that was/were stroked synchronously, while the pointing data only supported single hand ownership. The authors conclude that both body image and body schema can accommodate a fake limb, but only the body image supports multiple fake limbs. Finally, Folegatti et al. [9] search for clarification of multiple fake limb ownership in the body's spatial constraints. Using two rubber right hands placed to the left and even more to the left of the real right hand (both real hands out of sight), they show that, when one rubber hand is stimulated synchronously and the other remains unstimulated, ownership can occur over the synchronously stimulated one, regardless of how far it was from the real hand, according to questionnaire ratings and proprioceptive drift measures. Furthermore, when both hands are stimulated, only the closest rubber hand can be owned, but the drift data did not support this. The authors conclude that we should consider spaces that 'belong to the body', 'can belong to the body', and 'can affect the body' in a new conceptual framework in the context of ownership.

There are many possibilities as to why the aforementioned studies seem so contrasting, and here we only highlight a few. For example, there is no single correct way to measure the sense of embodiment (see [14] for an overview). Although subjective measurement through questionnaires seem straightforward and many studies use a variation of the one by Botvinick and Cohen, the exact questions are often slight variations of the originals, and some authors only choose to use a select group of the questions rather than all. Moreover, there is no agreement on how to interpret these results: some authors consider any high responses sufficient to say an illusion took place, while others solely rely on a significant difference even though all results are at null-level or even at a negative/below mid-line level. Regarding objective measures, although drift was used in the original experiment, many studies have since then shown that proprioceptive drift may in fact not be an indicator of ownership [1]. At the same time, there is no rule for what the best alternative is. Typical alternative physiological measures include SCR, heartbeat deceleration, and skin temperature, and for each it is difficult to argue whether in fact ownership is being measured, or whether there are no confounding factors. In the general use of SCR, for example, not only do personal characteristics determine the character of one's skin conductance overall, but an experimenter must also always be aware of the possibility of a habituation effect occurring. For heart rate deceleration, it has been shown that deceleration is greater for stimulus processing without a stressor than for response preparation with a stressor [5], meaning that although it is used in RHI studies as a response to a threat, the stressor is in fact not key to the deceleration. Similar to the case of drift, skin temperature has been shown to be unrelated to ownership, and changes as a result of the experimental setup [22]. Additionally, some previous studies would require reanalysis using adequate statistical methods before one can draw conclusions on body ownership.

This discussion illustrates the difficulties that occur when designing an RHI experiment, and in many cases these are valid reasons as to why one's results differ from previous studies. We remark that the goal of this paper is not to describe the 'perfect' SHI experiment; we aim to provide more insights and clarity on the topic. Luckily, there are a few key findings that almost all studies agree upon: they argue that a combination of top-down and bottom-up processing mechanisms must lie central to the RHI, rather than solely bottom-up processing as suggested by Botvinick and Cohen. It is further indisputable that *some* participants experienced *something* in the direction of embodiment regarding an extra limb (except in the study by Armel and Ramachandran), although it is not always agreed that this experience is in fact ownership.

1.2 Rubber Hand Illusions in Augmented Reality

Although many RHI studies take place in reality or VR, comparably there seems to be less interest in investigating the AR case. One of the few studies is the one performed by IJsselsteijn et al. [12], where the RHI is performed in three forms of media, all with synchronous brush stroking: reality, 'VR' (virtual 2D projections of rubber left hand and brush), and 'mixed reality' (MR; virtual 2D projection of rubber left hand, real brush). We remark that despite the names, both the 'VR' and 'MR' could both be considered variations in AR, since the rest of the setting (e.g., table and right arm) remain real. One of the main goals of the study was to explore how well the RHI can be reproduced in various media. Questionnaire and drift results showed that ownership occurred in all cases, albeit weaker in the mediated conditions, likely due to the 2D representation of the rubber hand. Moreover, according to the questionnaire, ownership in the VR case was stronger than in the MR case. The authors explain that this could be due to inconsistencies in the MR case as a result of the not so seamless integration of real and virtual.

Suzuki et al. [27] studied the interaction of exteroceptive with interoceptive processes in a 'cardiac RHI' using an AR setting. The authors captured depth and color information of the real left hand, which was out of view. With this, participants could see a 3D virtual projection of their real hand, on which redness of the skin changed either synchronously or asynchronously with the participant's heartbeat, or did not change but was combined with synchronous/asynchronous stroking with a brush. The results showed that, in all cases with synchronous cardiac feedback, the participants subjectively reported a sense of ownership. Moreover, when participants were able to move the virtual hand, ownership also occurred in the asynchronous cardiac feedback case. Still, the synchronous tactile case without any cardiac feedback, resulted in the highest reported ownership, even though the virtual hand became 'broken' (or discontinuous) when the virtual brush appeared. The authors conclude that multisensory integration of exteroceptive and interoceptive signals can modulate body ownership, but suspect that movement signals dominate the influence of the interoceptive signals. More important, they suggest that *a priori* willingness to accept the virtual hand could have played a role, while simultaneously being influenced by the interoceptive sensitivity of participants. In a more recent study, Feuchtner and Müller [8] investigated how users could control far away real objects using a virtual arm in AR. They examined four types of visualizations: a virtual arm without a real hand (removed in the style of diminished reality), a disconnected virtual hand without a real hand, an abstract virtual hand without a real hand, and a virtual arm with a real hand. The latter condition could be considered a supernumerary hand illusion setup. Their results showed that ownership only occurred in the condition where the virtual arm was shown without the real hand; other conditions, including the supernumerary hand condition, showed negative ownership results. They concluded that sufficient hand realism and connectedness preserve ownership in AR, whereas seeing the real hand disrupts the illusion.

Finally, some studies have investigated embodiment scenarios in augmented virtuality (AV), where head-mounted displays are used rather than CAVE VR to show the real body in a virtual environment. In one particular study, Jung et al. [13] investigated the influence of virtual body representation on object size estimation using conditions comparing an AV scenario with a personalized real hand image to a VR scenario with a generic hand model. They showed the personalized real hand image increased both subjective ownership of the hand and spatial presence compared to a generic virtual hand model, and furthermore supported participants in correctly estimating the size of a virtual object in the proximity of their hand.

It is strange that, although AR clearly triggers the RHI differently than VR, there are so few studies of it in AR compared to VR. One crucial reason is that VR and AR are at quite different stages of development: while consumer VR technology is already quite common, most AR developments are still in research phase. At a theoretical level, for example, it is common to consider a completely computer-generated environment as VR and the real world with virtual elements as AR. However, using Milgram and Kishino's definition of virtual objects [19], the study by Suzuki et al. [27] would no longer be considered AR because the virtual arm is in fact not 'virtual' but 'mediated' [23], even though it is quite agreed that this experiment is indeed in AR. At a more practical level, real-time seamless integration of real and virtual elements is still very challenging, while it is still not clear what levels (or even what kind of) fidelity is required for different experiences, as demonstrated by IJsselsteijn et al. [12].

Despite the extra challenges AR adds to this already complex area of research, both of the above studies have demonstrated why RHI research in AR is necessary. Specifically, the study by IJsselsteijn et al. shows that the RHI is not experienced equally across various media, meaning that there is a crucial difference in the way we step into these experiences. One of these differences is already expressed by Suzuki et al.: *a priori* willingness to accept a virtual limb, compared to *a priori* plausibility of accepting a rubber hand in the classic RHI. This suggests different ways of dealing with different types of non-realness: in the classic RHI, the rubber hand is not real, but clearly exists; in AR, the virtual hand is not real and it does not exist. One open question in particular then remains: what roles do the realism of the virtual hand and the participant's mindset due to the choice of medium play when interacting with virtual limbs? To illustrate, IJsselsteijn et al. used different visualizations of the rubber hand to compare the different media, and found that 'correct' appearance of the fake limb may

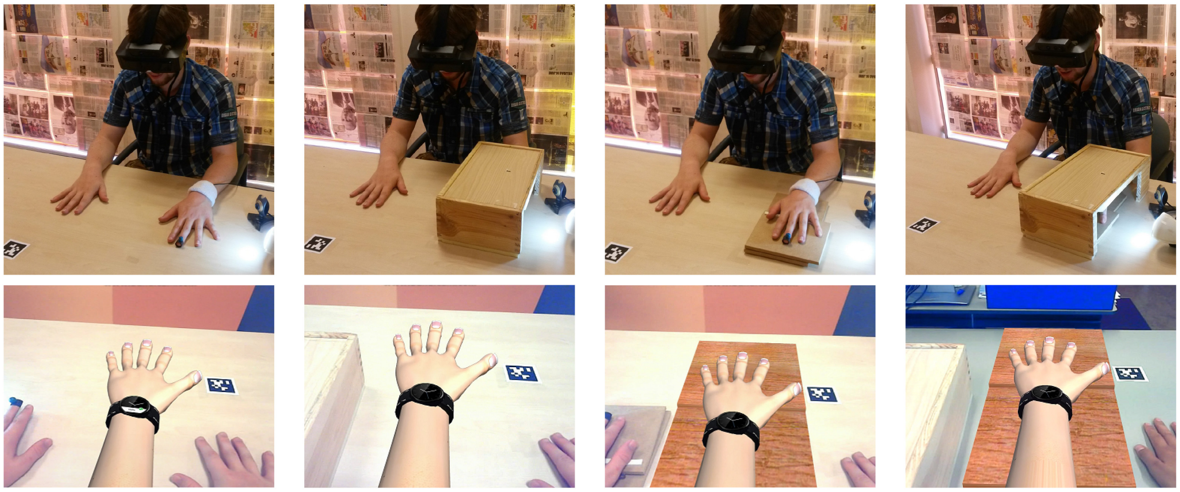


Fig. 1. Setup of the pilot experiment for various conditions, from [10]. A vibration motor is attached to the real left wrist, and the virtual left hand is wearing a smartwatch. A wooden box was used to hide the real hand in order to compare 'single' and supernumerary virtual arm conditions. (Top) front view, (bottom) participant's view; (left) visible real lefthand with smartwatch notification and active/no movement, (left-center) real left hand not visible and active/no movement, (right-center) visible real left hand and passive movement, (right) real left hand not visible and passive movement.

lead to more ownership. This is reflected in the study by Suzuki et al. where a 3D projection of the real hand consistently evoked positive levels of ownership. In contrast, typical RHI studies in VR use a virtual hand/body model that clearly intend to represent human hands/bodies, but fail to pass as 'realistic', while still evoking high levels of ownership.

1.3 Objective

The goal of this article is to study the feasibility of the SHI, and to study the required perceptual-motoric conditions for this illusion in AR, specifically. That is, the research question we investigate with this study is: What levels of ownership, agency, and arm-location are experienced over a supernumerary virtual arm while the real arms are still visually present in AR? We have elaborated on the current challenges in the area of SHIs in reality and RHIs in AR, and propose to approach both areas through novel 'augmented reality supernumerary hand illusion' (ARSHI) experiments. We do this in two steps, a pilot experiment and a main experiment, where the former served as an initial feasibility study. The contributions of this study include insights into perceptual and motoric factors that affect the sense of ownership, agency and arm-location in the ARSHI, and discussions on considerations that must be made when measuring and evaluating RHIs in general.

2 PILOT EXPERIMENT

The goal of this pilot study was to take a first step into understanding the fundamental criteria for an SHI in AR. Here we give a concise description of the pilot experiment; an elaborate description can be found in [10]. Twenty-seven participants experienced a certain multimodal stimulation for 3 minutes, which was always a combination of different levels of visible number of hands, visual-tactile synchrony and visual-motor synchrony. After this, they filled out a questionnaire on arm ownership, agency, and arm-location; see Figure 1 for the setup of the experiment. Here we shall only discuss two methodological factors in particular: the appearance of the virtual hand, and the visual-tactile stimulation by virtual smartwatch.

The virtual hand model was based on similar hand models that had been previously used in various RHI studies in VR [26]. Although there have been indications that corporeality of the rubber/virtual hand could play a role in the illusion, for example, when comparing a human-like hand with an abstract block [28], there was little support in the literature for the necessity of a more complex or higher quality virtual hand model in AR, therefore a corporeal hand shape with skin texture was deemed sufficient for this pilot, see the bottom row of Figure 1. Regarding the visual-tactile stimulation, IJsselsteijn et al. [12] mentioned various issues that affected ownership in both mediated conditions as opposed to the unmediated condition, one of which being sensory conflict after brushing a flat surface instead of an arm. To overcome such conflicts in this pilot experiment, visual-tactile stimulation was implemented by adding a virtual smartwatch, which would flash with a notification screen, and synchronously/asynchronously vibrate. Early iterations of this pilot study showed an improvement in compellingness of the visual-tactile stimulation when using a smartwatch compared to a simple tapping mechanism.

Regarding the results, overall the reported ownership remained negative (below midpoint) on a 7-point Likert scale. We suspect that the participants were possibly not inclined to accept the virtual hand as a real hand due to its model-like appearance and the obvious visual dissonance between virtual and real. The latter was moreover possibly emphasized by the clear difference between real and virtual in the environment as a whole, since the results are lower than expected compared to VR studies. This nonacceptance of the hand could be a result of the a priori willingness factor mentioned in Section 1.2. However, this is just speculation: it is unclear whether the results would have been different if the participants had been warned beforehand that the virtual arm looked so virtual, but to still try to pretend as much as possible that it was real. To overcome the hand-fakeness barrier, we decided to continue the main experiment with a projection of the participant's real hand, similar to that of [27], as to provide the virtual hand with a link to reality.

3 MAIN EXPERIMENT

As stated in Section 1.3, the goal of this study was to examine both the feasibility of the SHI and the requirements for this illusion in AR. In the pilot experiment, we found that hand appearance could be such a requirement, although it was not necessarily due to corporeality as indicated by related work, but could additionally be associated with overcoming the fakeness barrier between reality and virtuality. In the main experiment, we examine this possibility by using a projection of the participant's real hand as the virtual hand, and furthermore examine the induction of the SHI by means of visual-tactile and visual-motor stimulation, which are the most common forms of stimulation in other RHI studies.

3.1 Method

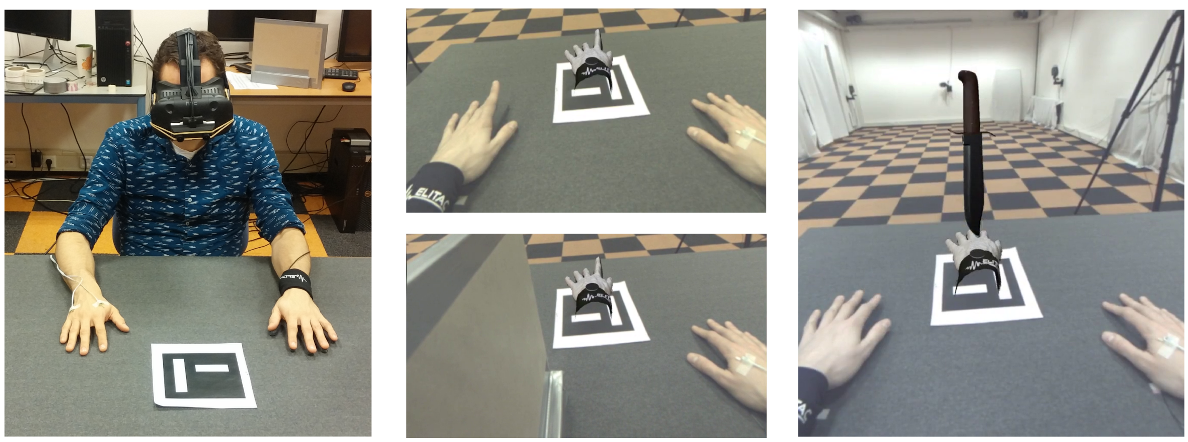
Six conditions were used in this experiment, which consisted of specific combinations of different levels of visible number of hands, visual-tactile synchrony and visual-motor synchrony, see Table 1. It was expected that the number of hands would affect the level of arm-location, not agency, and importantly also not the level of ownership, showing that ownership over a supernumerary hand is as feasible as a 'single' hand illusion. Visual-tactile synchrony would affect the level of ownership, but not agency or arm-location, and visual-motor synchrony would affect both the levels of ownership and agency, but not arm-location. These six specific combinations of factors were chosen rather than all possible combinations, in order to better inspect the feasibility of the SHI, as opposed to examining the relations between the various types of stimulation. Specifically, *C1* and *C2* compare the single virtual hand illusion with the SHI, while the other conditions compare necessary synchronous stimulation for the SHI to asynchronous control stimulation. Remaining combinations, such as the those with both asynchronous visual-tactile and asynchronous visual-motor stimulation, were excluded because they would not provide more insight on SHI feasibility than the included control conditions.

A detached virtual left hand was added approximately in front of the participant, and in one condition a small wall was used to block the view of the real left hand, see Figure 2. Visual-tactile stimulation consisted

Table 1. The Stimulation Used in the Six Conditions of the Main Experiment

Condition	*Hands*	*Visual-tactile*	*Visual-motor*
C1	real and virtual	synchronous	synchronous
C2	virtual	synchronous	synchronous
C3	real and virtual	synchronous	-
C4	real and virtual	asynchronous	-
C5	real and virtual	-	synchronous
C6	real and virtual	-	asynchronous

C2 is a single virtual lefthand condition where the real right hand was still visible, whereas all others are variations of supernumerary hand conditions where all three real and virtual hands are visible. *C3* and *C4* did not include any hand movements, and *C5* and *C6* did not include vibrotactile smartwatch notifications.



(a) Front view of experiment during certain condition.

(b) Participant's view of experiment during *C1* (top) and *C2* (bottom).

(c) Participant's view of experiment during the threat during *C1*.

Fig. 2. Experiment setup. The participants were seated at a table with both arms extended in front of them. The left wrist was covered by an elastic band containing one vibration motor. The electrodes for the SCR reading were attached to the right hand. During stimulation participants would see a virtual version of their left hand located above the marker.

of a simplified virtual smartwatch face that would flash white for 250 ms and vibrate for 250 ms to resemble a notification, which would either be synchronous or asynchronous. There was a standard inter-notification onset interval of 2 seconds, and in the asynchronous case both the visual and tactile stimuli received a random offset between −500 ms and +500 ms, meaning in some occasions the visual part came first, and in others the tactile part came first. The visual-motor stimulation was always active movement in the form of a tapping motion of the left index finger, and also occurred synchronously or asynchronously, where the latter was implemented as a delay of 1 second. In the cases where there was no movement (*C3* and *C4*), the participant was instructed to keep their real hand still, and in the cases with movement (*C1*, *C2*, *C5*, *C6*) the participant was instructed to make their own lifting movements with their finger, and specifically to not copy the virtual hand.

For subjective measurement of ownership, agency, and arm-location, a questionnaire was used, see Table 2. This questionnaire consisted of nine questions that were answered on a 7-point Likert scale ranging from "strongly disagree" to "strongly agree". For objective measurement of ownership, skin conductance was recorded before and after a threat to the virtual hand, in the form of a virtual knife making stabbing motions towards the

Table 2. The Post-Condition Questionnaire

Q	*Question*	*Measure*
Q1	It seemed as if the virtual hand was my hand.	Ownership
Q2	It seemed as if I had three hands.	Ownership
Q3	It seemed as if (1) the vibrations I felt were caused by the notification of the virtual smartwatch, and (2) the movements I felt were caused by the movements of the virtual fingers.	Ownership
Q4	It seemed as if the real hand were slowly becoming digital.	Control
Q5	It seemed as if I was controlling the virtual hand.	Agency
Q6	It seemed as if the virtual hand was controlling my will.	Control
Q7	It seemed as if my left hand was at two different locations.	Location
Q8	It seemed as if I felt the vibrations and movements at the location of the virtual hand.	Location
Q9	It seemed as if I felt the vibrations and movements somewhere between the real and virtual hand.	Location

The questions were presented in a random order, and answered on a scale from "strongly disagree" to "strongly agree". Depending on the condition, the formulation of Q3, Q8 and Q9 were changed to concern either only 'vibrations' or 'movements'. Ownership is measured through three main questions that each reflect a different aspect of ownership: attribution, changing the body image, and the source of experienced sensations [11, 14], and the location questions cover three possible directions of change: multiple locations, full shift, and partial shift. Questions Q4 and Q6 were control questions.

virtual hand, without it actually touching the surface of the virtual hand. SCR was chosen over other objective measures in order to allow comparison of results with other SHI studies.

3.2 Material

A video see-through head-mounted display (HMD) was made by mounting an OvrVision Pro onto an HTC Vive, which has a 1080 by 1200 per eye resolution. The device was calibrated using the standard OvrVision Calibration tool. The application was developed in Unity 5.5.0 using SteamVR. The scripts were written in C# using Microsoft Visual Studio 2015. The participant's left hand was recorded using a Kinect for Windows, and was displayed as if on top of a marker, see Figure 2(b). The distance between the center of the real hand and the center of the marker was approximately 25cm horizontally (such that the virtual hand would appear in between the two real hands horizontally) and 25cm in depth. This way, both of the real hands and the virtual hand were in the field of view without requiring head movement. The participant wore an Elitac Tactile Display with a single vibrotactile motor fastened to the wrist by a black elastic band. The left hand was raised by approximately 8mm using a cardboard hand cut-out covered in the same cloth that covered the table, in order to bypass the limited depth resolution of the Kinect. Using a BioSemi package, electrodes were attached to the right hand of the participant using a conductive paste. See Figure 2(a) for the front view of the participant. One computer, which could only be seen by the experimenter, ran ActiView to monitor the SCR, and another computer ran Unity. Because of problems that would occur after building, the experiment was run in play mode in Unity Editor, leading to a refresh rate of approximately 35 Hz. For general AR applications, this rate is not ideal, and in our case was most noticeable during head movements. Since the participants were instructed to only focus on the virtual hand and the real hands were always in view, eliminating the need to move the head at all, we expect this hardly affected the experimental outcomes.

3.3 Participants

Thirty participants were recruited from the student pool of the university through advertisement by email and during lectures, with notification of a raffle for three 25-euro gift cards to an online department store as incentive.

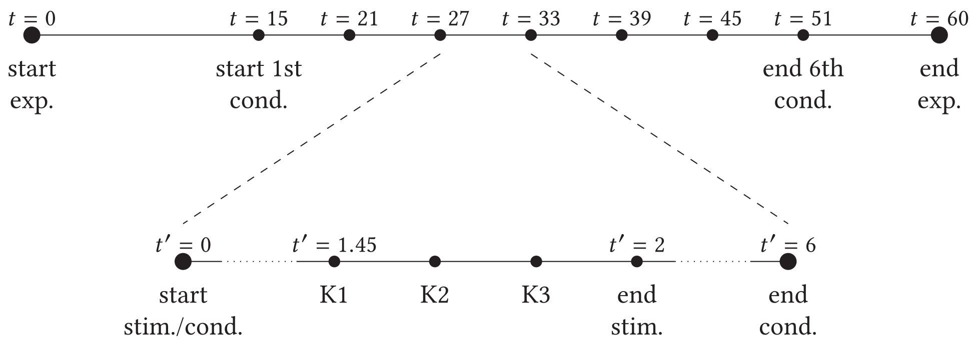


Fig. 3. Outline of procedure for a single participant: (top) outline entire experiment, (bottom) outline single condition. t indicates the time within the entire experiment, t' the time within a single condition; both are listed in minutes. K1 indicates the moment the virtual knife appears floating above the virtual hand, K2 the moment it starts to make a stabbing motion, and K3 the moment it disappears (while the virtual hand remains visible).

We remark that there was no student-teacher relation between the students and the experimenter/advertiser, therefore they were aware that not participating would not affect their course or education program progress. The average age was 22.7 (range 17–29, standard deviation 2.8), 24 male, 6 female. Twenty participants stated that they knew about or had experience with AR, and 10 stated that they knew about or had experience with body illusions. It was confirmed that none of these factors influenced the results of this experiment.

3.4 Procedure

Participants were sent an information letter by email, including a copy of the consent form, at least a week prior to participation. This information letter included a verbal description of the threat they could expect during the experiment. Upon arrival in the laboratory, each participant was asked to read and sign the consent letter a final time, to wash and dry their hands, and take a seat at a table. Then the experimenter applied the electrodes to the medial phalanges of the index and middle fingers of the right hand. If for some reason this did not provide a signal or the participant had cuts or callouses on the fingers, the thenar and hypothenar eminences of the same hand were used. Next, the experimenter attached the vibrotactile motor to the left wrist, and measured the inter-pupilar distance (IPD), which in turn was adjusted in the HMD. The experimenter then put the HMD on the participant, and asked whether the participant could see the environment correctly; the IPD was then adjusted if necessary. Once the participant found this was the case, then the first session was started.

During each session, the participant was asked to experience a certain type of stimulation for 2 minutes (see Table 1). After 1 minute and 45 seconds of stimulation, a knife would appear and float above the hand, while stimulation continued. After another 5 seconds, this knife would move in a stabbing motion towards the hand, and after another 5 seconds, the knife would disappear and only the virtual hand remained. After a final 5 seconds, the virtual hand would also disappear and the session was over. After the session, the participant would remove the HMD and fill out a questionnaire on a tablet where the questions were presented in random order (see Table 2). After approximately 4 minutes, the next session was started. Each session, including the break, would therefore take approximately 6 minutes.

After all six conditions were completed, which occurred in a unique random order for each participant, the participant was asked a few post-experiment questions to check for lasting effects of the experiment. These were based on those used by Kilteni et al. [15]. Finally, the purpose of the different conditions was explained to the participants and they were allowed to ask any questions they had. They were also asked whether they would like to take part in the raffle. An outline of the portion of the experiment that took part in the laboratory can be found in Figure 3.

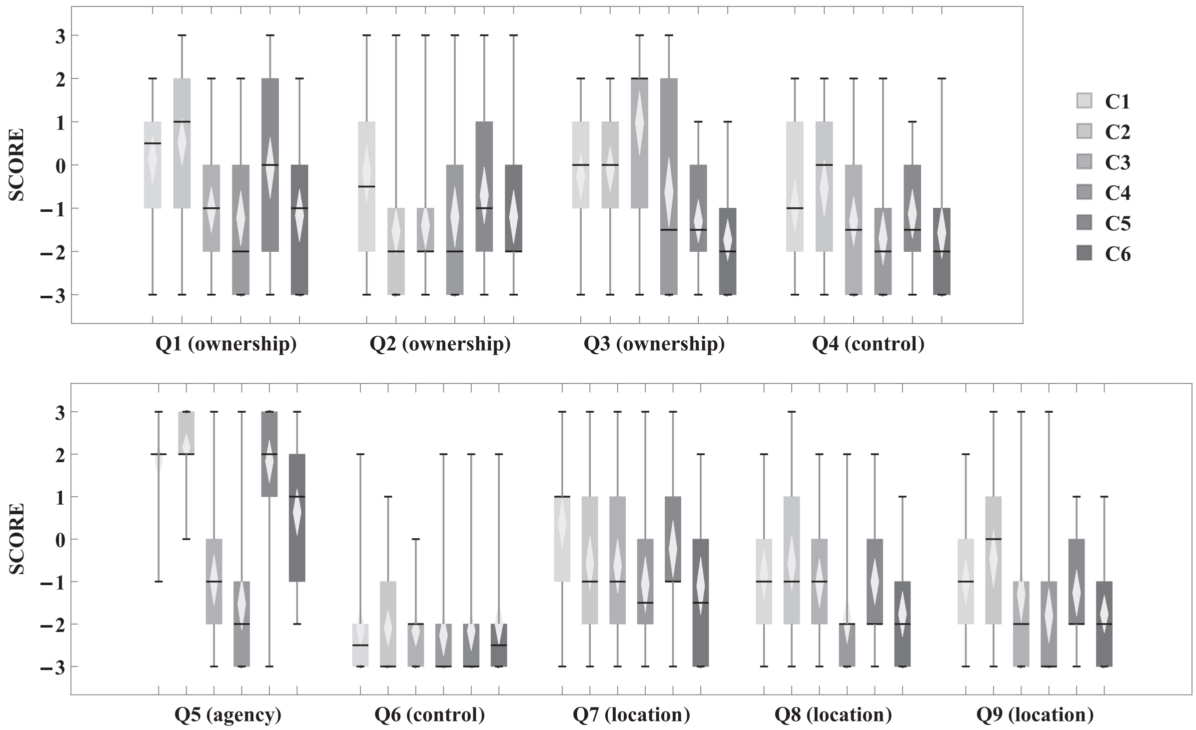


Fig. 4. Boxplots of all questionnaire responses, on a seven-point Likert scale where −3 corresponds to "strongly disagree" and +3 to "strongly agree". The medians are marked with a black horizontal line, and the means +/- mean standard errors are marked as light grey diamonds.

At least 3 days after participation, each participant was called and asked a second round of post-experiment questions. In two cases where the participants could not be reached, the questions were asked by email. There were no lasting effects for any participants. The medical ethics committee of the local university hospital did not raise objections to the execution of this experiment.

4 RESULTS

4.1 Questionnaire Data

First, all questionnaire results were converted to numbers for simplicity, where "strongly disagree" corresponded to −3, and "strongly agree" to +3. Then, for each question a Friedman test was used to test for an overall effect over conditions in SPSS. All questions except for Q6 resulted in a significant effect, and post-hoc analyses with Bonferroni correction were run. See Figure 4 and Table 3 for the results. Generally, the responses for *C1* (real and virtual left hands, synchronous VT, synchronous VM) and *C2* (only virtual left hand, synchronous VT, synchronous VM) were highest, and *C4* (real and virtual left hands, asynchronous VT, no VM), and *C6* (real and virtual left hands, no VT, asynchronous VM) lowest in comparison.

4.2 Participant Comments

Of the 180 possible post-condition comments and the 60 possible post-experiment comments, there were a total of 69 and 48 non-empty meaningful comments, respectively. Each of these comments was labeled, where each comment could receive more than one of 19 labels, and the labels ranged from 'immersive experience' ("…the

Table 3. The Results of the Friedman Tests per Questions

Q	$\chi^2(5)$	p	post-hocs at $p < 0.00333$
Q1	62.827	<0.0001	*C1,C2,C5* > *C3,C4,C6* with exception *C3* = *C5*
Q2	19.435	0.002	*C1* > *C2,C3*
Q3	48.653	<0.0001	*C1,C2* > C5,*C6* and *C3* > *C1,C4,C5,C6*
Q4	25.265	0.0001	*C1,C2* > *C4*
Q5	102.327	<0.0001	*C1,C2* > *C3,C4,C6* and *C5,C6* > *C3,C4* and *C5* > *C6*
Q6	3.383	0.641	-
Q7	23.582	0.0003	*C1* > *C4,C6*
Q8	34.694	<0.0001	*C1,C2,C3* > *C4*
Q9	24.482	0.0002	*C2* > *C6*

All Friedman test results are asymptotic, and all accompanying post hoc Wilcoxon signed rank test results with Bonferroni correction are exact.

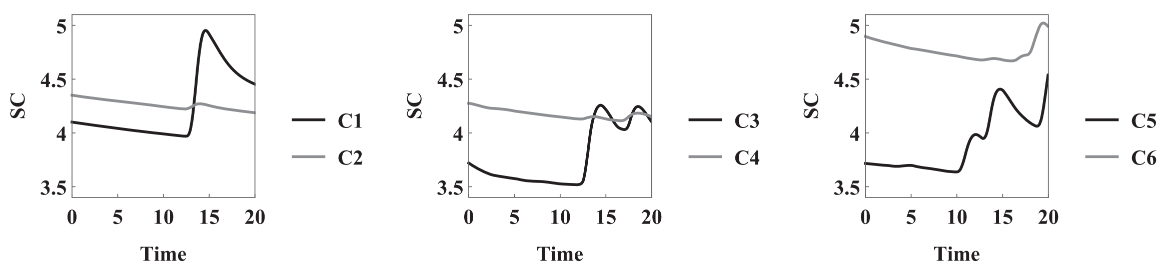
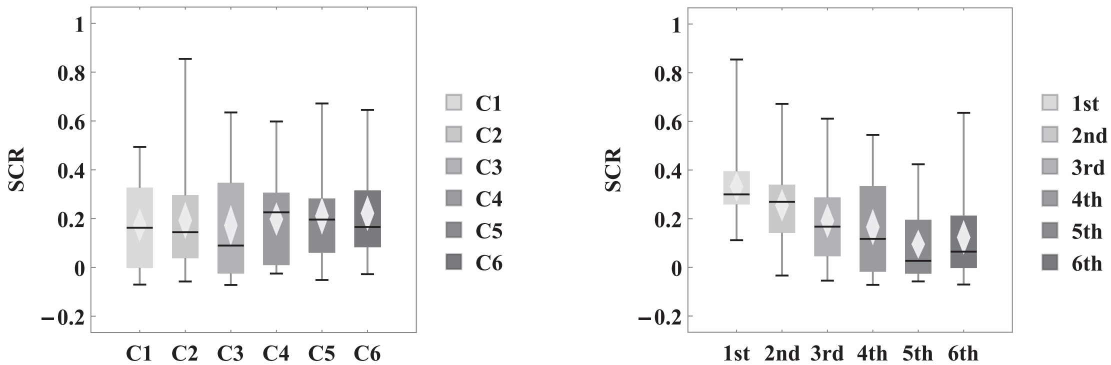


Fig. 5. Representative skin conductance plots of one participant for each condition. The vertical axis represents skin conductance (SC) in microSiemens, the horizontal axis time in seconds. The single SCR value used for analysis was computed by taking the average of the first 10 seconds of the SC curve (baseline) and subtracting this from the maximum level that occurs in the last 10 seconds of the SC curve (knife presence), with finally a log transformation log(microSiemens +1).

hand felt completely real") to 'meta experiment' ("What about a real knife instead of a virtual knife?"). Notable cases were that 14 participants remarked that the hand did not look realistic enough in a post-condition comment, 27 participants remarked positively regarding the experiment in general in a post-experiment comment, 10 participants remarked experiencing some form of discomfort in a post-experiment comment, and 8 participants remarked an experienced imbalance over modalities in a post-condition comment.

4.3 SCR Data

4.3.1 Initial SCR Data Analysis. In order to inspect only phasic changes in skin conductance, rather than applying a high pass filter, the SCR values were computed using a difference function: take the average of the 10 seconds before the threat appeared and subtract this from the maximum level that occurs in the 10 seconds that the knife is present (5 seconds static knife presence, and 5 seconds stabbing motion); see Figure 5 for an example of skin conductance during these 20 seconds. A Friedman test over all the SCR data showed there was no significant effect of condition, $\chi^2(5) = 2.636; p = 0.756$. However, using a post-hoc Friedman test on the same SCR values, but now arranged by order of performed condition, there was a significant habituation effect towards the threat, $\chi^2(5) = 38.621; p < 0.0000003$ (see Figure 6). Using multiple Wilcoxon Signed Ranks Tests with Bonferroni correction, the differences were that the *first* presented condition was significantly greater than the *third*, *fourth*, *fifth* and *sixth* (all $p < 0.0003$), and *second* was significantly greater than *fifth* and *sixth* (all $p < 0.001$).



(a) SCR by condition. (b) SCR by performed order.

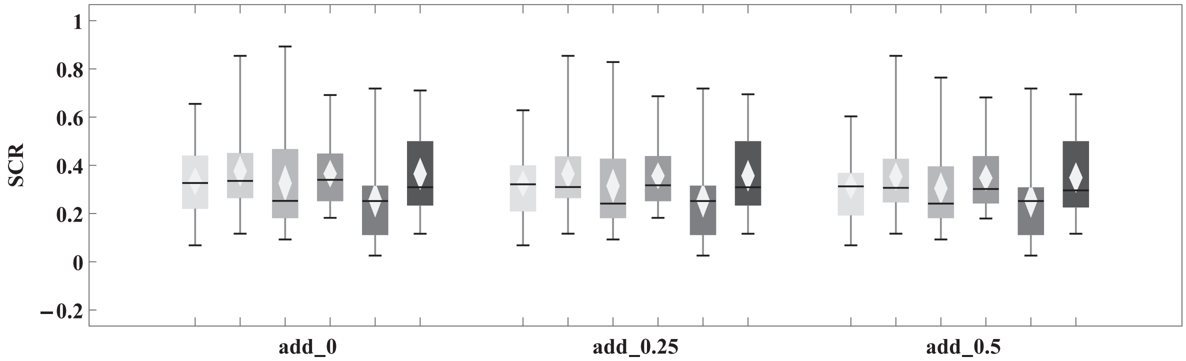
Fig. 6. Boxplots of the SCR values. Each value is computed by subtracting the average of the 10 seconds before the appearance of the knife from the maximum value in the 10 seconds of knife presence, then adding 1 and taking the log with base 10, i.e., log (microSiemens + 1). The medians are marked with a black horizontal line, and the mean +/− mean standard error are marked as light grey diamonds.

Table 4. Results of the Linear Regressions for Each Condition

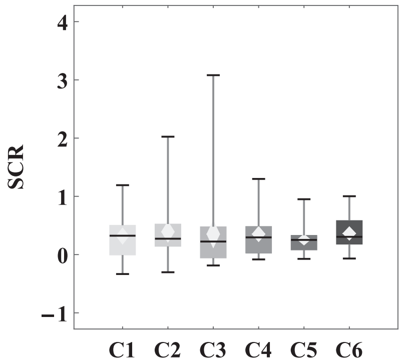
Function	*Fit*	R^2
*C1*_SCR = 0.334 − 0.147 ∗ log(case)	$F_{1,28} = 10.015; p = 0.004$	0.263
*C2*_SCR = 0.376 − 0.170 ∗ log(case)	$F_{1,28} = 10.062; p = 0.003$	0.276
*C3*_SCR = 0.325 − 0.144 ∗ log(case)	$F_{1,28} = 6.092; p = 0.020$	0.179
*C4*_SCR = 0.366 − 0.158 ∗ log(case)	$F_{1,28} = 12.529; p = 0.001$	0.309
*C5*_SCR = 0.257 − 0.048 ∗ log(case)	$F_{1,28} = 0.674; p = 0.419$	0.024
*C6*_SCR = 0.364 − 0.120 ∗ log(case)	$F_{1,28} = 4.591; p = 0.041$	0.141

Each regression was constructed by using 30 data points, where the SCR values were a function of the natural log of the case number (i.e. first → log(1), second → log(2), etc.). All but *C5* lead to a reliable linear regression.

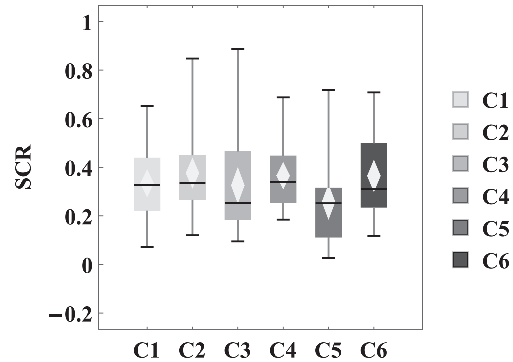
4.3.2 Habituation Correction. According to Montagu [20], the 'true' SCR reaction corrected for habituation should be equal to the intercept of the linear regression, where SCR is a function of the natural log of the case number. In the following, we discuss steps that were taken to investigate the transformed-to-true data. Since each condition consisted of different stimuli, a linear regression could not be performed on the entire SCR data, but at the same time, each participant had too few data points to perform reliable linear regressions. Therefore, we performed one linear regression for each condition, see Table 4. Three basic transformation approaches were then examined: multiplicative, additive, and rotational; see Figure 7. For the additive method, we included the option of SCR personalization. Specifically, we transformed the SCR value by adding a value that depends on the regression slope and log of the case number. If this value exceeds the maximum SCR value for this participant, then the transformed SCR value is reduced by subtracting a proportion c of this difference. We examined values $c = 0, \frac{1}{4}, \frac{1}{2}$ and performed a Friedman test for each value. For each case, there was a significant effect between conditions, $\chi^2(5) = 18.362, p = 0.003$, and in each case post hoc Wilcoxon Signed Rank Tests with Bonferroni correction showed that *C5* was significantly less than *C2*, *C4,* and *C6* (all $p < 0.001$). Unfortunately, this leads to effects that negates basic assumptions, such as that *C5* would generally be greater than *C6*. We do not expect an alternative approaches to personalization to improve this matter. A Friedman test on the results with a multiplicative transformation without personalization showed no significant effect, $\chi^2(5) = 3.886, p = 0.566$. For the rotational approach, the SCR values are vertically translated such that the regression line intersects the origin,



(a) SCR, additive transformations.



(b) SCR, multiplicative transformation.



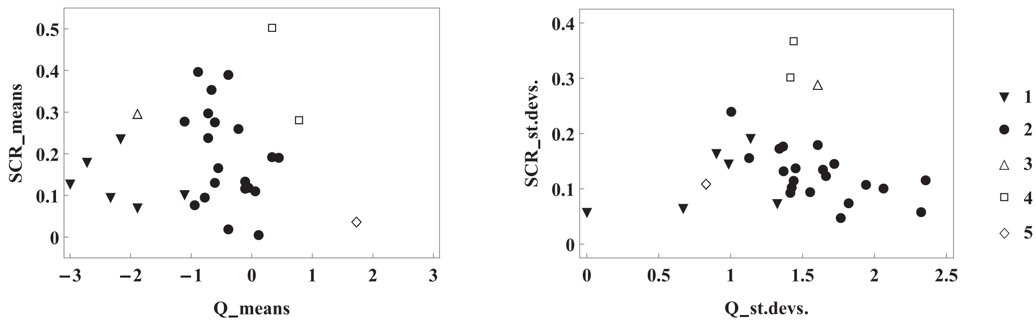
(c) SCR, rotational transformation.

Fig. 7. Boxplots of the SCR values with additive, multiplicative and rotational transformation, in log(microSiemens + 1). Note that the vertical axis in the multiplicative case has a different scale. The medians are marked with a black horizontal line, and the mean +/− mean standard error are marked as light grey diamonds.

then rotated by the angle between the regression and the horizontal axis, and finally translated back. A Friedman test showed a significant effect of condition, $\chi^2(5) = 18.686, p = 0.002$, and again the post hoc Wilcoxon Signed Rank Tests with Bonferroni correction showed that *C5* was significantly less than *C2*, *C4,* and *C6* (all $p < 0.001$), as in the additive cases. In summary, none of the habituation correction approaches led to more insight into the differences that were found in the questionnaire data.

4.4 Relationship between Questionnaire and SCR Data

Since the questionnaire data regarding ownership *did* result in significant effect while the SCR data did *not*, it is important to evaluate how participants responded *individually* post-hoc. To do this, we compare each participant's SCR 'character' from the data without any habituation correction to their questionnaire response 'character'. The SCR character is defined here as the range in which a participant's SCR lies; similarly, the questionnaire response character is defined as the range in which the responses to Q1, Q2, and Q3 lie, which are the three positively formulated non-control questions regarding ownership. We can approach these ranges by performing a cluster analysis on the means and standard deviations of both questionnaire and SCR data, resulting in an analysis of a 4-dimensional space. Considering the number of data points, hierarchical clustering was chosen for the analysis on the four variables: SCR_means, SCR_st.devs., Q_means, and Q_st.devs. Clustering with between-groups



(a) Clustering shown on means.

(b) Clustering shown on standard deviations.

Fig. 8. Cluster graphs after clustering on four variables: questionnaire response mean and standard deviation over Q1, Q2, and Q3 over all conditions, and SCR mean and standard deviation over all conditions. Two graphs are used to display the five resulting clusters on all four variables, and we emphasize that both graphs represent the same clustering result.

average linkage, squared Euclidean distance, Z-score normalization, and by applying the stopping rule in the agglomeration schedule, we find the clustering as shown in Figure 8. We see that, besides the four possible outlying points (cluster numbers 3, 4, and 5), there are two large clusters: the participants with low questionnaire response means and low standard deviations who also have low to medium SCR means (cluster number 1), and those with high questionnaire response means and high standard deviations who also have low to high SCR means (cluster number 2). Executing the analysis for the remaining three variables, i.e., all but SCR_st.devs., provides the same clustering. A multiple correlation test with SCR_means as dependent variable and Q_means and Q_st.devs. as independent variables does not result in a significant correlation, $F(2, 29) = 0.160$, $p = 0.853$ with $R^2 = 0.012$.

5 DISCUSSION

The questionnaire results of the main experiment demonstrate that the SHI is indeed feasible in AR. Regarding the ownership questions, the conditions with multiple synchronous stimulation (i.e., both visual-tactile and visual-motor) resulted in a positive experience of attribution (Q1), and asynchronous conditions led to a negative feeling of attribution. Unexpectedly, both single synchronous conditions (i.e., either visual-tactile only or visual-motor only) were experienced as weak. Curiously, only the multiple synchronous stimulation case with three visible hands led to a moderate experience of having three hands (Q2), while all other conditions, even those with three visible hands, did not cause this experience. This feeling was most negative for the case with only two visible hands, indicating a higher weighting to visuals in the context of changing the overall body image throughout the experiment. Regarding the source of the sensations (Q3), synchronous visual-tactile stimulation alone showed the highest response, which is surprising, considering the weak attribution results for this condition. The participant's comments on imbalance over modalities often concerned the multiple synchronous stimulation conditions, where it felt as if the visual-tactile stimulation had more of an impact than the visual-motor stimulation, but this is not reflected in the results of the ownership questions. Had this been the case, then the multiple synchronous stimulation conditions would have equal responses as the single synchronous visual-tactile condition. All conditions showed negative experiences of the real hand becoming digital (Q4; control). We remark that although this question was intended as a derivation of the original "turning rubbery" question by Botvinick and Cohen [3], the "turning digital" version could have been interpreted differently. While both terms could reflect a materialistic aspect, 'digital' could also reflect a visual-only aspect, namely whether the arm seemed to be made up of pixels. Since a video see-through HMD was used here, participants could have been inclined to respond less negative than the rubbery counterpart in a classic RHI; however, we do not expect that this occurred since the results do not differ strongly from those of the synchronous visual-tactile condition reported by Botvinick and Cohen.

Both multiple stimulation conditions resulted in the highest sense of agency (Q5), and the single synchronous visual-motor stimulation followed closely. Approximately half of the participants also reported a positive sense of agency in the asynchronous visual-motor condition, which is consistent considering they still had control, albeit delayed. As expected, the single visual-tactile conditions resulted in negative feelings of agency. Also, the control question of whether the virtual arm was in control (Q6) resulted in negative responses in all conditions, as expected. For the experience of arm-location, most participants reported feeling multiple left hand locations in the multiple synchronous stimulation condition where three hands were visible (Q7), and this sensation was negative for most participants in all other cases. Regarding the question of whether the sensations were felt at the location of the virtual hand (Q8), all results were negative; however, the conditions with at least synchronous visual-tactile stimulation were least negative. Finally, regarding whether the sensations were felt between the real and virtual hand (Q9), we see a weak response for the condition with only two visible arms; all other conditions were negative. Together these arm-location results indicate that when there is multiple synchronous stimulation, if two left hands are in view then two left-hand positions are experienced, as opposed to an experience of drift when there is only one left hand in view.

5.1 Consistency with Other Research Findings

These questionnaire results concur with related work considerably. With respect to SHIs for example, we achieved similar response levels to a variety of questions used in Guterstam et al.'s [11] experiments 1 (compare to *C3* and *C4*) and 5 (compare to *C1* and *C2*, with the exception of synchronous visual-motor stimulation). For the RHI in AR, again similar response levels are found in Suzuki et al.'s [27] tactile condition (compare to *C2*, with the exception of synchronous visual-motor stimulation) and IJsselsteijn et al.'s VR condition (compare to *C2*, with the exception of synchronous visual-motor stimulation) [12]. Notably, higher ownership results were found in our study compared to the disconnected virtual hand and visible real hand conditions in Feuchtner and Müller's study [8] (compare to *C1* and *C2*, with the exception of synchronous visual-tactile stimulation).

Although further comparison with more studies from Sections 1.1 and 1.2 is hardly possible due to large differences in experimental setup, the conclusions are similar: it is possible to experience a supernumerary limb as one's own alongside the real hand in AR. Nevertheless, a few remarks on methodological differences between these studies and our study must be made. First, it should be noted that these studies focused only on the concept of ownership, and their questions therefore intend to measure ownership only. Q8 and Q9 are examples of questions that are regarded as ownership questions in other experiments, but in our experiment are considered arm-location questions because they explicitly ask about the location of certain experienced sensations. The study by Guterstam et al. regards Q9 as an ownership control question, indicating this may indeed not measure ownership in the first place [11]. Second, agency was not measured in any of these studies, and is a contribution of our study to the existing body of literature. Specifically, according to the questionnaire results of the main experiment, active synchronous visual-motor stimulation led to a sense of agency over the virtual hand, while active asynchronous stimulation did not, regardless of the visual presence of the real hand, as expected. Third, although we have discussed visual-tactile stimulation as a general approach, our pilot and main studies used a virtual smartwatch rather than a more common (virtual) brush stroking. Since the responses in our and other studies are similar, we do not expect that this difference affected the experience. There is also no reason to believe that brushing should be considered 'more correct' than a smartwatch; Botvinick and Cohen [3] did not explain or defend their initial choice for brushing. By contrast, the limited movement freedom may have influenced the ownership results: the subjective ownership results for the movement condition in [27] are slightly higher than the results of our movement condition without tactile feedback, indicating that in fact more ownership could have been experienced with more movement freedom. Although this does not influence the major conclusions of our study, namely that the SHI is feasible in AR, it does provide a fruitful area for future research.

5.2 Visual Realism and Willing Suspension of Disbelief

The differences in results between the pilot and main experiment suggest that hand appearance played a crucial role in the experience of ownership. The virtual hand model used in the pilot experiment was insufficient, while a real hand projection in the main experiment yielded positive subjective results. The fact that the model was insufficient was surprising, since it did not differ strongly from virtual hand models typically used in VR RHI studies. Moreover, although the hand projection *was* adequate in the main experiment, there were still many comments on the bad hand appearance, such as that the virtual hand “looked a bit deformed”, “was not super realistic”, and “looked weird.” Three participants specifically noted the greyish skin tone of the virtual hand, which likely became apparent due to color shifts in the display. Even though it has been shown that skin color can affect the experience of ownership in [16], this study also showed that both extreme skin tones elicited an illusion, thus we do not expect this to have influenced our study.

One participant explained that although the virtual hand was obviously a projection of their own real hand, it was not *really* their hand. These contrasts show that it is not yet clear whether improvement of the sense of ownership should *only* be sought in the improvement of realism; in AR, it could be more a matter of passing a barrier of belief that is emphasized by the dissonance between reality and virtuality, which is in turn possibly determined by *a priori* willingness to accept virtual limbs. This indicates a new top-down factor at play in the sense of embodiment. Although willingness is an uncommon concept in RHI literature, it has been mentioned in the context of general media use. For example, by drawing upon phenomenological similarities between VR and narrative text, Ryan [24] discusses the necessity of Coleridge’s willing suspension of disbelief in order to regard non-actual possible worlds as actual. Translating Coleridge’s views to VR and AR, he argues that if the creator of the virtual/mixed environment ensures that the environment is as truthful as possible, then the user would willingly accept the fake environment as real [7]. That is, the user actively decides to accept the fiction if it seems plausible, which differs greatly from, for example, the ideas on how presence passively occurs: without choice as a result of, for example, system variables [17]. Still, willing suspension of disbelief may be related to immersive tendencies, and in turn presence, in the sense that an individual with high immersive tendencies may have those personality traits that make them more prone to decide to accept fiction [29]. Future research should further investigate this relationship.

5.3 Questionnaire versus SCR

It is puzzling why the questionnaire results show such clear differences, while these same differences are absent in the SCR data. The SCR results also do not seem to reproduce similar SCR results from other studies, although these results did not always use the same calculation, namely the maximum value in the threat window minus the average in the pre-threat window: Armel and Ramachandran [2] seem to only use the maximum value in the 5-second threat window , while Guterstam et al. [11] use maximum minus minimum, both within the 5 second threat window. For reference, both methods were also compared on our data, but still did not find the expected differences (Friedman test $\chi^2(5) = 15.695; p = 0.008$, *post hoc* with Bonferonni correction $C4 < C6$ $p = 0.002$ two-tailed, and Friedman test $\chi^2(5) = 5.232; p = 0.388$, respectively). We suspect that the habituation effect that occurred in this experiment was stronger than any possible effect based on condition. At the same time, there was not enough data to correct for individual habituation effects, therefore only group level effects could be inspected. It is noteworthy that, although SCR habituation is a very basic, common, and strong effect, it is, to our knowledge, not mentioned in any RHI studies that use it as an objective measure. We suggest that future studies critically evaluate habituation effects. We remark that there is so far no evidence for a similar habituation effect in the questionnaire data. By defining questionnaire habituation as less variation in the responses over time, it can be examined by comparing the standard deviations of all questionnaire responses split by order. The absence of a habituation effect in the questionnaire responses was subsequently confirmed, $F(5, 145) = 1.699$; $p = 0.139$.

To relate the questionnaire and SCR data, a cluster analysis was executed. We found two main groups among the participants: a group with overall low but consistent questionnaire responses with low to medium SCRs, and a group with overall high and more inconsistent questionnaire responses with low to high SCRs. We elaborate why the choice of cluster analysis over, for example, linear regression, is justified. Prior to this experiment, one could have reasonably hypothesized that participants with low SCR responses would also respond more negatively to the questionnaires (and analogous for high and positively). The results of the main experiment, however, showed that participants did not react in this way: the significant differences that appeared in the questionnaire data did not also appear in the SCR data. To find out whether certain links could be deduced from the different data types, a linear regression would have been insufficient, since it is clear that the questionnaire and SCR data would not lie on the hypothesized line. Instead, we were inclined to search for possible groups of participants that defined regions in the 4-dimensional search space. The multiple correlation test was furthermore not significant; however, this is not surprising since the questionnaire mean and standard deviation seem to predict the range within which the SCR mean occurs, not necessarily the SCR value itself.

We emphasize that a correlation between the multiple data types is not at all self-evident. That is, in general, it is not necessarily the case that when questionnaire results show the same significant differences as the SCR results (which was not the case in our study), that these differences occurred for all participants. Whether similar results occurred for all participants over multiple data types must therefore always be further inspected or verified with a suitable test or analysis. This is currently lacking in the literature, which contributes to the seemingly contrasting results as explained in Section 1.1. Admittedly, verification has become more difficult considering exclusion of specific participants is commonplace. In some cases, exclusion is based on an initial filtering stage as a result of certain participants simply not being prone to the illusion, such as in the study by IJsselsteijn et al. [12]. In other cases, this occurs after too many zero-response occurrences in SCR data for a single participant, such as in the study by Armel and Ramachandran [2], while moreover the threshold for 'too many' is arbitrarily chosen. A consequence of participant exclusion becomes clear from the cluster analysis performed in this study: those participants in cluster number 1 with no subjective proneness to the illusion could still show medium SCRs, while some participants in cluster number 2 that would be considered SCR zero-responders elsewhere could still show high questionnaire responses. In other words, the idea behind exclusion of these participants is to remove those that do not show the necessary response, while in fact the participants are possibly capable of response in a different valid measure. We encourage future RHI studies to critically evaluate the need for exclusion, and consider suitable tests and analyses to further inspect the results of multiple data types.

6 CONCLUSION

The goal of this article was to examine the feasibility of the SHI, together with the requirements for this illusion in AR. We did this by measuring ownership, agency, and arm-location in a variety of conditions that considered the visible number of hands, visual-tactile stimulation, and visual-motor stimulation as factors. Through a pilot study, we found that our virtual hand model, which was similar to those typically used in VR RHI studies, was insufficient to produce a sense of ownership. In the main experiment, we examined the SHI with a projection of the participant's real arm, and measured the sense of embodiment through questionnaire and SCR. The questionnaire results showed that participants subjectively experienced a sense of ownership and agency in both single and supernumerary hand conditions. Also, the results suggested that subjective arm-location depended on the number of visible hands: two left hands led to an experience of an additional left-hand location, whereas one left hand resulted in a shift in experienced left-hand location. Unfortunately, the SCRs did not confirm an objective experience of ownership, likely due to a strong habituation effect, and we recommend the consideration of habituation in future RHI studies using SCR. However, by linking questionnaire and SCR data, we found that the participant pool was characterized by two groups, which in turn demonstrated that illusion proneness can be more complex than initially expected. Together, these results demonstrate the feasibility of the SHI in AR, where hand appearance plays a crucial role, but is in turn possibly related to a willingness to accept a virtual limb. We propose future research to consider willingness as a new top-down factor in RHI studies performed in AR, which may be determined by personality traits, and to further examine the intricacies involved in creating appropriate virtual limbs beyond corporeality and basic realism.

REFERENCES

[1] Zakaryah Abdulkarim and H. Henrik Ehrsson. 2016. No causal link between changes in hand position sense and feeling of limb ownership in the rubber hand illusion. *Attention, Perception, & Psychophysics* 78, 2 (2016), 707–720. DOI:https://doi.org/10.3758/s13414-015-1016-0

[2] K. Carrie Armel and Vilayanur S. Ramachandran. 2003. Projecting sensations to external objects: Evidence from skin conductance response. *Proceedings of the Royal Society of London B: Biological Sciences* 270, 1523 (2003), 1499–1506. DOI:https://doi.org/10.1098/rspb.2003.2364

[3] Matthew Botvinick and Jonathan Cohen. 1998. Rubber hands 'feel' touch that eyes see. *Nature* 391 (1998), 756. DOI:https://doi.org/10.1038/35784

[4] Wen-Yeo Chen, Hsu-Chia Huang, Yen-Tung Lee, and Caleb Liang. 2018. Body ownership and the four-hand illusion. *Scientific Reports* 8, 1 (2018), 2153. DOI:https://doi.org/10.1038/s41598-018-19662-x

[5] V. De Pascalis, R. J. Barry, and A. Sparita. 1995. Decelerative changes in heart rate during recognition of visual stimuli: Effects of psychological stress. *International Journal of Psychophysiology* 20, 1 (1995), 21–31. DOI:https://doi.org/10.1016/0167-8760(95)00023-L

[6] H. Henrik Ehrsson. 2009. How many arms make a pair? Perceptual illusion of having an additional limb. *Perception* 38, 2 (2009), 310–312. DOI:https://doi.org/10.1068/p6304

[7] Anthony J. Ferri. 2007. *Willing Suspension of Disbelief: Poetic Faith in Film*. Lexington Books, Lanham, MD.

[8] Tiare Feuchtner and Jörg Müller. 2017. Extending the body for interaction with reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 5145–5157.

[9] Alessia Folegatti, Alessandro Farnè, R. Salemme, and Frédérique de Vignemont. 2012. The rubber hand illusion: Two's a company, but three's a crowd. *Consciousness and Cognition* 21, 2 (2012), 799–812. DOI:https://doi.org/10.1016/j.concog.2012.02.008

[10] Carolien Gilbers. 2017. *The Sense of Embodiment in Augmented Reality: A Third Hand Illusion*. Master's thesis. Utrecht University.

[11] Arvid Guterstam, Valeria I. Petkova, and H. Henrik Ehrsson. 2011. The illusion of owning a third arm. *PloS ONE* 6, 2 (2011), e17208. DOI:https://doi.org/10.1371/journal.pone.0017208

[12] Wijnand A. IJsselsteijn, Yvonne A. W. de Kort, and Antal Haans. 2006. Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence: Teleoperators and Virtual Environments* 15, 4 (2006), 455–464. DOI:https://doi.org/10.1162/pres.15.4.455

[13] Sungchul Jung, Gerd Bruder, Pamela J. Wisniewski, Christian Sandor, and Charles E. Hughes. 2018. Over my hand: Using a personalized hand in VR to improve object size estimation, body ownership, and presence. In *Proceedings of the Symposium on Spatial User Interaction*. ACM, 60–68.

[14] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 21, 4 (2012), 373–387. DOI:https://doi.org/10.1162/PRES_a_00124

[15] Konstantina Kilteni, Jean-Marie Normand, Maria V. Sanchez-Vives, and Mel Slater. 2012. Extending body space in immersive virtual reality: A very long arm illusion. *PloS ONE* 7, 7 (2012), e40867. DOI:https://doi.org/10.1371/journal.pone.0040867

[16] Marilia Lira, Julia H. Egito, Patricia A. Dall'Agnol, David M. Amodio, Óscar F. Gonçalves, and Paulo S. Boggio. 2017. The influence of skin colour on the experience of ownership in the rubber hand illusion. *Scientific Reports* 7, 1 (2017), 15745. DOI:https://doi.org/10.1038/s41598-017-16137-3

[17] Matthew Lombard and Theresa Ditton. 1997. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication* 3, 2 (1997), JCMC321. DOI:https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

[18] Matthew R. Longo, Friederike Schüür, Marjolein P. M. Kammers, Manos Tsakiris, and Patrick Haggard. 2008. What is embodiment? A psychometric approach. *Cognition* 107, 3 (2008), 978–998. DOI:https://doi.org/10.1016/j.cognition.2007.12.004

[19] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems* 77, 12 (1994), 1321–1329.

[20] J. D. Montagu. 1963. Habituation of the psycho-galvanic reflex during serial tests. *Journal of Psychosomatic Research* 7, 3 (1963), 199–214. DOI:https://doi.org/10.1016/0022-3999(63)90004-7

[21] Roger Newport, Rachel Pearce, and Catherine Preston. 2010. Fake hands in action: Embodiment and control of supernumerary limbs. *Experimental Brain Research* 204, 3 (2010), 385–395. DOI:https://doi.org/10.1007/s00221-009-2104-y

[22] Marieke Rohde, Andrew Wold, Hans-Otto Karnath, and Marc O. Ernst. 2013. The human touch: Skin temperature during the rubber hand illusion in manual and automated stroking procedures. *PloS One* 8, 11 (2013), e80688. DOI:https://doi.org/10.1371/journal.pone.0080688

[23] Nina Rosa, Peter Werkhoven, and Wolfgang Hürst. 2016. (Re-)examination of multimodal augmented reality. In *Proceedings of the 2016 Workshop on Multimodal Virtual and Augmented Reality (MVAR'16)*. ACM, New York, Article 2, 5 pages. DOI: https://doi.org/10.1145/3001959.3001961

[24] Marie-Laure Ryan. 1999. Immersion vs. interactivity: Virtual reality and literary theory. *SubStance* 28, 2 (1999), 110–137. DOI: https://doi.org/10.1353/sub.1999.0015

[25] Michael Schaefer, Hans-Jochen Heinze, and Michael Rotte. 2009. My third arm: Shifts in topography of the somatosensory homunculus predict feeling of an artificial supernumerary arm. *Human Brain Mapping* 30, 5 (2009), 1413–1420. DOI: https://doi.org/10.1002/hbm.20609

[26] Mel Slater, Daniel Pérez-Marcos, H. Henrik Ehrsson, and Maria V. Sanchez-Vives. 2009. Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience* 3, 2 (2009), 29. DOI: https://doi.org/10.3389/neuro.01.029.2009

[27] Keisuke Suzuki, Sarah N. Garfinkel, Hugo D. Critchley, and Anil K. Seth. 2013. Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia* 51, 13 (2013), 2909–2917. DOI: https://doi.org/10.1016/j.neuropsychologia.2013.08.014

[28] Manos Tsakiris, Lewis Carpenter, Dafydd James, and Aikaterini Fotopoulou. 2010. Hands only illusion: Multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Experimental Brain Research* 204, 3 (2010), 343–352. DOI: https://doi.org/10.1007/s00221-009-2039-3

[29] David Weibel, Bartholomäus Wissmath, and Fred W. Mast. 2010. Immersion in mediated environments: The role of personality traits. *Cyberpsychology, Behavior, and Social Networking* 13, 3 (2010), 251–256. DOI: https://doi.org/10.1089/cyber.2009.0171

Received June 2018; revised March 2019; accepted May 2019