Embodiment and telepresence : toward a comprehensive theoretical framework

Citation for published version (APA):

Haans, A., & IJsselsteijn, W. A. (2012). Embodiment and telepresence : toward a comprehensive theoretical framework. *Interacting with Computers*, 24(4), 211-218. <https://doi.org/10.1016/j.intcom.2012.04.010>

DOI:

[10.1016/j.intcom.2012.04.010](https://doi.org/10.1016/j.intcom.2012.04.010)

Document status and date:

Published: 01/01/2012

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Embodiment and telepresence: Toward a comprehensive theoretical framework[☆]

Antal Haans^{*}, Wijnand A. Ijsselsteinj

Eindhoven University of Technology, The Netherlands

ARTICLE INFO

Article history:

Available online 9 May 2012

Keywords:

Media technology

Presence

Embodiment

Body schema

Body image

Sensory-motor integration

ABSTRACT

What explains the experience of “being there” in a simulated or mediated environment? In recent years, research has pointed to various technological and psychological factors deemed important in eliciting this so-called experience of telepresence, including interactivity, sensory-motor integration, media transparency, and distal attribution. However, few theories exist that can combine these findings in a coherent framework. In the present paper, we formulate such a theoretical framework. We will argue that the experience of telepresence is a consequence of the way in which we are embodied, and that it extends naturally from the same ability that allows us to adjust to a slippery surface, or to the weight of a hammer. The importance of embodiment in the understanding of telepresence has been stated before, but these works have not yet fully addressed what it means to be embodied. We argue that “having a human body” means having a specific morphology, a body schema, and a body image. Subsequently we describe how tools and technological artifacts may be incorporated at each of these levels of embodiment, and the implications thereof for the experience of telepresence.

© 2012 British Informatics Society Limited. All rights reserved.

1. Introduction

Media technologies allow us to meet and converse with colleagues overseas, to operate in hazardous environments from a safe location, and to visit a new house even before it has been built. What all these technologies have in common is that they replace or augment people's immediate surroundings by digital content. This content may be simulated by a computer (in case of virtual reality), or may originate from recording devices located at a remote site (in case of teleconferencing and teleoperation systems). When this augmentation is successful, the users of these technologies may start to feel as if they are physically present in the mediated or simulated environment. Such experience of “being there” might even include a sense of ownership over the hands of the slave robot in a teleoperation setup or over your avatar's body parts in virtual reality (e.g., Haans and Ijsselsteinj, 2007). Cole et al. (2000) describe such an experience when they used a teleoperation system at Johnson Space Center in Houston: “... *there is a misidentification of the sense of ownership of one's own body, this being transferred into a set of steel rods and stubby robotic hands with little visual similarity to human arms.*” (p. 167).

Minsky (1980) coined the term *telepresence* when describing the phenomenon of being there as experienced by operators of teleoperation systems. Others have suggested a distinction between telepresence and (virtual) presence, where the latter refers to the experience of being physically present in a simulated rather than a mediated remote environment (e.g., Sheridan, 1992). However, since both telepresence and virtual presence are expected to be similar in their psychological underpinnings, a distinction between the two types of enabling media technologies is unnecessary; particularly so, in the context of the present framework (cf. Ijsselsteinj et al., 2000). In this paper, we will use the term telepresence to mean both: The experience of being there in a mediated or simulated environment.

Since the early 1990s, researchers from various disciplines have started to investigate the technological and psychological determinants of telepresence as well as its effects on users of immersive media technologies. According to Steuer (1992), telepresence is determined by two technological factors: Interactivity and vividness. Interactivity refers to the extent to which users can alter the stream of sensory information or make modifications to the virtual or mediated environment (see also Sheridan, 1992). Interactivity, however, should not be confused with engagement or involvement. The experience of telepresence and the experience of being involved with whatever occurs in the virtual environment are two logically distinct phenomena. As Slater (2003) argued, presence is about form not symbolic content. One will feel unmistakably present in the perfect virtual reality simulation, but it will be as emotionally captivating, scary, or boring an experience as being in the real life environment upon which it was modeled.

[☆] This paper has been recommended for acceptance by D. Murray.

^{*} Corresponding author. Address: Human-Technology Interaction group, School of Innovation Sciences, Eindhoven University of Technology (IPO 1.35), P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands. Tel.: +31 40 2475237; fax: +31 40 2449875.

E-mail address: A.Haans@tue.nl (A. Haans).

Although logically distinct, telepresence and emotional involvement may be empirically related (Slater, 2004). Length and weight are logically distinct attributes, but they are positively correlated in the population of human bodies. Similarly, telepresence and involvement are found to be correlated in at least some virtual environments (e.g., Riva et al., 2007; Gorini et al., 2011). A captivating content or narrative may affect telepresence, for example, by diverting one's attention from those technological imperfections that otherwise would impede the experience.

The second technological factor that Steuer identified is vividness, which refers to the level of detail in the digital environment as well as to the number of sensory modalities that are addressed (Steuer, 1992). Note that vividness does not necessarily entail pictorial realism. Regardless of how genuine or fake the environment may look, moving your head while wearing a head-mounted display and seeing the virtual world being updated in a corresponding and consistent manner is already a strong cue to elicit telepresence (Dinh et al., 1999). Loomis (1992) explains this observation through the process of distal attribution. Any physical object in the environment distinguishes itself from creations of the mind by a set of invariant, or lawful, sensorimotor relationships. For example, when turning your head the image on the retina moves in the opposite direction with a matching displacement. When the mediating technology allows for these sensorimotor contingencies to be registered, or enacted, one will experience the digital content as having a specific physicality in time and space.

Other researchers have pointed to the importance of media transparency (Ijsselstein et al., 2004; Lombard and Ditton, 1997). According to these authors, telepresence occurs when the user forgets about the technology, and thus feels and acts as if the mediating technology itself were not there. This requires that the media technology is seamlessly integrated in the action–perception loop. Lombard and Ditton (1997) even define telepresence as the perceptual illusion on non-mediation. This definition expands the experience of telepresence to other technologies including hearing aids and eyeglasses.

Despite considerable advances in our understanding of telepresence, few theories exist that can combine this knowledge in a coherent framework. In the present paper, we formulate such a theoretical framework. We will argue that the experience of telepresence is a consequence of the way in which we are embodied, and that the capability to feel as if one is actually there in a technologically mediated or simulated environment is a natural consequence of the same ability that allows us to adjust to, for example, a slippery surface or the weight of a hammer. The importance of embodiment in the understanding of telepresence has been stated before (Biocca, 1997; Ijsselstein, 2005; Slater and Usoh, 1994). Biocca (1997), for example, argues that the body itself is not much different from the technologies we use: Both mediate the communication between the mind and the physical world. Similarly, Ijsselstein (2005) argues that the capability of the central nervous system to incorporate tools and technological artifacts into our body representations is a crucial mechanism behind the experience of telepresence. These authors, however, have not yet fully addressed what it means to be embodied. Different tools, for example, appear to be incorporated into our embodiment in different ways (Haans and Ijsselstein, 2007).

Based on scholarly works by Metzinger (2003, 2006) and Gallagher (1986, 2005a,b), we propose that humans are embodied on three different levels: first, the level of body morphology, secondly, the level of the body schema, and thirdly, the level of the body image. In the following sections we describe each of the three levels of embodiment in more detail. At the same time, we discuss how tools and technological artifacts may be incorporated at each

of these levels of embodiment, and the implications thereof for the experience of telepresence.

2. Embodiment and telepresence

In recent years, the term embodiment (or being embodied) has become popular in various disciplines of science and technology, including human–computer interaction (e.g., embodied interaction with computers (Dourish, 2001) and cognitive psychology (e.g., embodied cognition (Varela et al., 1991)). In these instances, the term embodiment is commonly used in the tradition of philosophers like Heidegger, Husserl or Merleau-Ponty: as being an active participant in the world (see e.g., Zahorik and Jenison, 1998). Embodied cognition, for example, postulates that it is through this participation, which is bounded by the characteristics and possibilities of the human body, that intelligence and the mind itself can be explained. Yet, how exactly are we embodied? According to Metzinger (2006), there are three different orders of embodiment, which he calls, first, second, and third order embodiment. These three orders of embodiment can be explained in terms of the morphology of the body, the body schema, and the body image. First order embodiment means having morphology only. Second order embodiment means having morphology and a body schema. Finally, third order embodiment means having all three: morphology, a body schema, and a body image.

2.1. Morphology

The human body has certain characteristics, including the number, kind, and location of limbs, muscles, and sensory receptors, that distinguish it from bodies of other animals, such as pigs, birds or bats. These morphological characteristics largely determine the animal's behavior. Having wings, for example, may, in general, be considered a necessary condition to fly. In other words, the morphological and physiological characteristics of the body both enable and constrain the animal's action possibilities. Scratching your own back when it itches can be annoyingly difficult, because the length and flexibility of the arm, as well as the degrees of freedom of its joints, do not allow you to reach the itching spot easily. At the same time, body morphology determines, to some extent, the quality of the individual animal's experiences. An eye, for example, is a prerequisite for sight. Since all human beings have a highly similar morphology, one can presume that others are capable of having the same or highly similar experiences. In contrast, if one lacks wings and a sense of echolocation, then it is hard to imagine what it would be like to be a bat (Nagel, 1974).

Having a body is the most fundamental way in which an organism can be embodied. Having morphology only constitutes what Metzinger (2006) calls first order embodiment. An example of first order embodiment would be a simple Braitenberg vehicle such as depicted in Fig. 1. This device consists of a mechanical body with wheels and two light sensors as eyes. Each light sensor is connected directly to the motor of the contra-lateral wheel. Since the two light sensors are located at a distance from each other, the automaton's direction of movement is dependent on how much light each sensor registers. If the rightmost sensor receives more light than the leftmost sensor, then the left wheel will spin faster than the right wheel. As a result the automaton will orient itself toward the light source. Moreover, the closer the automaton is to a light source, the more light each sensor receives, and thus the higher the speed of movement will be. For an outside observer it might appear as if the automaton has the intention to aggressively attack the light source. Note that this Braitenberg vehicle's

Fig. 1. A simple Braitenberg vehicle.

action possibilities are rigid rather than plastic: Being first order embodied, the automaton's action possibilities are hard-wired as morphological characteristics of its body.

2.1.1. Tools and morphology: physical extensions

We use tools and technological artifacts to extend, improve, or repair the perceptual and motor functions offered by our human morphology. Examples are plentiful: A hammer allows us to generate enough force to drive a nail into a piece of wood. A mechanical prosthesis can replace an accidentally lost limb. A magnifying glass allows us to see things otherwise too small for our own visual system. Similarly, we use media technology to interact with and operate in distant or simulated environments. However, a proficient use of these temporary extensions of our morphology, which includes the ability to rapidly switch between tools, requires not just a body (i.e., a morphology), but a dynamic body schema as well. In other words, fluent tool use requires an organism that is at least second order embodied.

2.2. Body schema

Our bodies are in a constant interaction with the environment, yet we generally do not pay much attention to what the body is doing. We can, for example, walk without having to consciously deliberate on every step we make. Consider a simple task as fetching a book from the top shelf of a bookcase. If the bookcase is not too high, then you “simply” extend your arm and grab the book in a single fluent motion. However, even such a simple task involves a complex pattern of muscle contractions which are required for extending the arm, standing on the tips of your toes, keeping balance, and grabbing the book, while at the same adjusting your posture for the weight of the book. It is due to the body schema that we can interact fluently with the environment despite the complexity of human motion. Since the body schema is involved in the continuous regulation of posture and movement, Gallagher (1986) defines the body schema as the “non-conscious performance of

the body” (p. 548), or as “a nonconscious system . . . of motor-sensory capacities that function below the threshold of awareness, and without the necessity of [conscious] perceptual monitoring” (Gallagher, 2005a, p. 234). The body schema necessarily constrains the action possibilities that the morphological characteristics of the body allow: There are many different ways in which we could fetch the book, but only one specific motion is endorsed at a given time. Another important aspect of the body schema is that it is dynamic rather than rigid (Gallagher, 2005a). It is due to this dynamic nature of the body schema that one can grab the book with the same effectiveness when, for example, one leg is injured or when holding a stack of books in the other hand. Similarly, the body schema can adjust itself to long term morphological changes, for example, due to growth or the accidental loss of a limb.

The most obvious function of the body schema is to keep track of the relative position of the body and its parts in time and space. According to Head and Holmes (1911) this involves an internal representation of the body, or a postural model of the body, which is constantly updated to account for the position and movement of the body and its limbs. They argue that there may be more than one single body representation, and discuss the possibility of another representation that maps a tactile sensation to a certain part of the body. However, determining the spatial position of body parts and the localization of tactile stimuli are not sufficient for the body schema to function. It requires, amongst others, a mapping of the space immediately around the body (e.g., to determine whether an object is within reach) and appropriate action selection (cf. Maravita et al., 2003). The body schema can, thus, best be described as a dynamic distributed network of procedures aimed at guiding behavior (cf. Kugel, 1969). Being involved in action selection, muscle actuation, and the positioning of the body and its parts in relation to objects in the world, this network of procedures combines the individual parts of our morphology into a coherent functional unity. In recent years, considerable progress has been made in uncovering how these body schema procedures are implemented in the central nervous system. This empirical evidence suggests that the body schema is constructed, and continuously updated, through sensorimotor integration. A more thorough discussion of this research falls outside the scope of the present paper, and is provided in various reviews (e.g., Maravita et al., 2003; Berlucchi and Aglioti, 1997; Holmes and Spence, 2004; Maravita and Iriki, 2004).

To an observer (including the owner of a body), it might appear that the body schema functions by means of a coherent whole-body representation (i.e., a homunculus of some sort), but such a representation is not required for the individual procedures to function (Minsky, 1988; Brooks, 1991). Although each individual procedure might serve rather simple tasks, such as determining whether something is happening around a certain part of the body, a network of procedures can, collectively, serve apparently complex tasks such as keeping balance. Furthermore, consistent with Gallagher's (1986, 2005b) conclusion that the body is anonymous at the level of the body schema, the network of procedures functions without conscious reference to the body and its limbs as owned by a person. The body schema, thus, requires neither a puppet, nor a puppeteer (and thus no Cartesian puppet-theatre either; cf. (Dennett, 1991).

2.2.1. Tools and the body schema: functional extensions

Head and Holmes (1911) already describe that objects in the environment can be incorporated into the body schema. Head's famous examples include the blind man's cane and the feather on a woman's hat. Regarding the latter example, one should understand that in the beginning of the 20th century it was fashionable for women to wear tall feathered hats (think of the famous 1910 painting “Lady with Black Feather Hat” by Gustav Klimt; see Fig. 2). Head

Fig. 2. Lady with Black Feather Hat. Oil painting by Gustav Klimt, 1910.

must have observed how elegantly and effortlessly these women passed through small doorways without disturbing the feathered piece of millinery on top of their heads. There is an increasing amount of empirical evidence to support such observations. In the following sections, we present four examples of research on the use of tools, such as sticks, rakes, and grabbers to increase our reaching space (or peripersonal space; e.g., (Holmes and Spence, 2004).

Yamamoto and Kitazawa (2001) delivered a stimulus to a finger of the left and right hand (or vice versa) of their participants, and asked them to judge the order in which the two tactile stimuli were presented by stretching the index finger of the hand they felt was stimulated first. With the arms uncrossed, people are generally quite proficient in making such temporal order judgments. However, people perform significantly worse when they make such judgments with the arms crossed (also Shore et al., 2002). Yamamoto and Kitazawa (2001) also investigated people's performance in making such temporal order judgments when the tactile stimuli were delivered to the tip of two sticks, which were held by the participant in each hand, rather than to the fingers. This time, temporal order judgments were made by using the sticks to press the levers on which the tip of the sticks were resting. They found that crossing the sticks had a similar detrimental effect on performance as crossing the arms, indicating that the sticks were incorporated into the body schema as extensions of the arms.

Further evidence is provided by research on bimodal neurons which respond to tactile stimulation of the hand as well as to visual stimuli near the hand (Maravita and Iriki, 2004; Iriki et al., 1996). Iriki and colleagues (1996) trained Japanese macaques in using a rake to retrieve a piece of food outside of the monkeys' normal reach. Using single cell recordings, they found that the monkey's bimodal neurons would not only respond to stimuli near the hand, but started to respond to stimuli near the rake as well. In other words, while the monkey retrieved pieces of food with the rake, the receptive field of its bimodal neurons had temporarily expanded to include the entire rake. No such expansion of the receptive field was found when the monkeys passively held the rake.

Berti and Frassinetti (2000) report on an experiment involving a patient P.P. who suffered from left-sided near space neglect after a stroke in the right hemisphere. The authors asked patient P.P. to perform a series of line bisection tasks, in which she had to indicate the midpoint of a line by means of a laser pointer. Consistent with near space neglect, she would perceive the midpoint of the line to

be closer toward the right when the line was located at a distance of 50 cm (in near space), but not when the line was located at a distance of 100 cm (in far space). However, when she had to point toward the midpoint of the line by means of a stick rather than a laser pointer, a displacement toward the right was observed for both far and near space. This illustrates that when using tools and rakes to touch objects that would otherwise be out of one's reach, the brain actually remaps the space around us to accommodate for the expansion of our reaching space. For patient P.P. this unfortunately included an expansion of her neglect as well.

Finally, Cardinali and colleagues (2009) demonstrated that the kinematics of manual grasping tasks is significantly altered after similar grasping tasks have been performed with a grabber. The demonstration of such after-effects offers additional evidence that tools are indeed incorporated in the body schema. Interestingly, the after-effects of using a grabber were not specific for grasping tasks but generalized to a pointing task as well. In a subsequent experiment, they asked blindfolded participants to indicate various positions on their trained right arm with their other hand. They found that the distance between the indicated position for the finger tips and the elbow of the right arm had increased after using the grabber. Taken together, the experiments indicate that representations of the arm, as part of the dynamic body schema, are functionally adaptive and become elongated when using tools that increase our reaching space.

These examples provide empirical evidence for the often made claim that tools and technological artifacts can be incorporated into the body schema. Or more correctly formulated, these examples demonstrate that the individual procedures in the distributed network that we call the body schema will adapt themselves to a tool. As a result, the tool becomes part, although temporarily, of the same functional unity as the various parts of our morphology. This, in turn, allows for a more fluent interaction with that tool. The components of immersive media technology, such as a head-mounted display, share a similar fate. Biocca and Rolland (1998), for example, investigated how well people would adapt to seeing the world through a head-mounted display and two cameras that were positioned slightly above and in front of their own eyes. The extent in which their participants adapted to the inter-sensory conflict caused by the head mounted display was determined by means of a pointing task. Although participants' performance on the pointing task had improved after a training session (i.e., performing pegboard tasks), they never reached their baseline level of performance (i.e., when not wearing the head-mounted display). That the participants' body schemas had adapted to the media technology became evident, however, from their performance on the pointing task after the head-mounted display was removed. Participants still made pointing errors, but this time in the opposite direction. Such visual-motor negative aftereffects indicate that the body schema has adapted itself to the new environment (e.g., Welch, 1978). These aftereffects are commonly observed with virtual reality applications (Groen and Werkhoven, 1998; Stanney et al., 1999) and might cause people, for example, to pour a beverage into their eyes when attempting to drink immediately after they step out of the virtual world (Strauss et al., 1995).

When the components of a virtual reality or teleoperation system are effectively integrated in the body schema, we can interact with the simulated or mediated environment as if the mediating technology were not there (i.e., the illusion of non-mediation). As a result, we feel as if we are physically there in the virtual or mediated environment. Whereas the adaptation of the body schema occurred almost instantaneously in the four examples above (Yamamoto and Kitazawa, 2001; Iriki et al., 1996; Berti and Frassinetti, 2000; Cardinali et al., 2009), the experiment by Biocca and Rolland (1998) illustrates that this is not always the case. Indeed, technological factors, such as the limited field of

views or delays, may constrain the adaptation of the body schema (also Held and Durlach, 1991; Welch et al., 1996). At the same time research on classic prism adaptation has demonstrated that active interaction with the environment is essential for the body schema to adapt to inter-sensory conflict (Held and Hein, 1958). In the experiment by Biocca and Rolland (1998) participants engaged in a series of pegboard tasks. Such tasks facilitate sensory recombination as they require many fine-grained movements to be executed. This reflects the importance of interactivity for the experience of telepresence as discussed in the introductory section of this paper (also Sheridan, 1992; Steuer, 1992). Welch and colleagues (1996), for example, found that people who could actively interact with the virtual environment experienced a stronger sense of being physically located in the simulated environment than individuals who were mere passive observers.

The experience of telepresence, however, entails not only the sense of being but might also include a sense of ownership over virtual body parts, or the arms and hands of a slave robot (e.g., as seen through a head-mounted display). Such experiences of ownership cannot be explained by incorporation into the body schema alone. Due to its anonymous nature, the incorporation of tools and technological artifacts into the body schema does not require that they are experienced as a part of the bodily self. We thus can make a distinction between functional extensions of the body (incorporation in the body schema) and phenomenological extensions of the self (Haans and IJsselstein, 2007; Gallagher and Cole, 1995). In the scheme of Metzinger (2006), this latter type of bodily extensions requires third order embodiment: Not just morphology and a body schema, but a body image as well. More specifically, third order embodiment requires the kind of higher-order consciousness that enables humans to hold a concept of their own body over time.

2.3. Body image

Gallagher (1986, 2005a,b) has defined the body image as our perceptions of the body, which include the way we see and experience our bodies, as well as any conceptual knowledge we have about our bodies. In contrast to the body schema, the body image is, in terms of Gallagher, not anonymous but owned. The body image can, in our point of view, best be described as a part of the process of consciousness. According to Edelman (2003, 2006) consciousness is the result of neural processes that allow for a large amount of refined discriminations and perceptual categorizations, by combining multimodal sensory information, and connecting such sensory information with memory content. These higher-order discriminations give rise to subjective experiences, for example the “redness” of red or “coldness” of cold, and include the unitary perceptual scene, emotions, and memories alike. If one accepts this formulation of consciousness, then the body image consists of those discriminations that pertain to the individual's own body (i.e., to those objects that the central nervous system has categorized as being a part of the physical body).

Although we have described the body image as a part of the process of consciousness, not all organisms are necessarily conscious of their body image. According to Edelman (2003, 2006), one can distinguish between two types of consciousness: primary and higher-order consciousness. All organisms that have primary consciousness have a body image. That is, they can make some minimal discrimination between their body and the environment out of all information that is available to them at a certain moment in time. Organisms with only primary consciousness, however, make such discrimination only for that brief period in time that Edelman calls “the remembered present”. They lack the higher-order consciousness that enables the linking of the remembered present with a remembered past, and an anticipated future. Organ-

isms with primary consciousness alone lack the capacity to be conscious about having a body image. Similarly, they lack the capability of being self-conscious. Put differently, with primary consciousness alone, an organism will possibly experience one bodily self after the other, without being able to combine these temporary bodily selves into a longer lasting conception of its own body.

In the scheme of Metzinger (2006), only organisms that have a body, a body schema and are conscious of having a body image are third-order embodied. Since this requires higher-order consciousness, only animals that are self-conscious (and, thus, in terms of Metzinger possess a phenomenological self model; see also (Metzinger, 2003, 2009) are third-order embodied. The traditional test to determine whether an animal is self conscious is the rouge test (Gallup, 1970). For this test, the animal is anesthetized and a patch of skin on the head or ear is marked with rouge (or another odorless dye). After the animal has fully recovered from anesthesia, a mirror is placed in the animal's cage. If the animal, after seeing the marks in the mirror, attempts to remove the marks from its body, then the inference is that the animal recognizes itself in the mirror and, thus, is self-conscious. Currently, only human beings, chimpanzees, bonobos, orang-utans, a single abnormally reared gorilla (Povinelli and Cant, 1995), and perhaps dolphins (Reiss and Marino, 2001), elephants (Plotnik et al., 2006) and magpies (Prior et al., 2008) have passed the rouge test, and can thus be expected to have higher-order body images.

2.3.1. Tools and the body image: phenomenological extensions

For a tool or technological artifact to become experienced as a phenomenological extension of the self, it is required that the central nervous system categorizes the object as a part of one's bodily self (as belonging within rather than outside the boundaries of the body). This process in which foreign objects are attributed to the self is again mainly dependent on the capability of the central nervous system to extract correlations between the various sensory modalities, upon which it reconstructs a meaningful representation of the world (and thus one's body; e.g., (Armel and Ramachandran, 2003). There is increasing evidence that infants, through their interaction with objects and other people (a process called body-babbling (Meltzoff and Moore, 1997), learn to distinguish between themselves and the environment by establishing body specific sensorimotor contingencies (Botvinick, 2004; Lackner, 1988; Rochat and Striano, 2000; O'Regan et al., 2005). Every event the infant perceives (e.g., the clapping of hands), whether self-initiated or not, consists of correlated multisensory impressions (e.g., the visual image and sound of clapping hands). In time, the infant learns that some of these patterns of sensorimotor contingencies are exclusively associated with the body, and hence self-specifying. Whenever a person exercises or perceives these sensorimotor contingencies, he or she “knows” (in a skill-like fashion; cf. (O'Regan et al., 2005) that the perceived object belongs to the body: When the visual image of clapping hands is accompanied immediately by a tactile sensation in the hands, then by inference it must be your hands that do the clapping.

By conforming closely to the patterns of self-specifying sensorimotor correlations, foreign objects, including tools and technological artifacts, can be incorporated in the body image. The ease with which such perceived body alterations can be induced is perhaps best illustrated by the rubber-hand illusion (Botvinick and Cohen, 1998) (for virtual versions of this illusion, see (IJsselstein et al., 2006; Slater et al., 2008). This body image illusion is induced by having a person watch a fake hand being stroked and tapped in precise synchrony with his or her own concealed hand by means of two brushes (see Fig. 3). After a few minutes of this kind of precise synchronous tapping and stroking of the fingers of the concealed hand and the fake hand, the person that is being stimulated in this way

Fig. 3. Experimental setup of the rubber-hand illusion.

might develop the vivid impression that the fake hand is actually his or her own.

In the rubber-hand illusion, there is a near perfect correlation between seen and felt stimulation. Moreover, this pattern of multi-sensory correlations matches the body-specific sensorimotor contingencies normally registered when you see and feel your own hand being stimulated by a brush. As a result, the central nervous system deduces that the fake hand must be part of one's own body. This is consistent with Richard Gregory's theory of perception as hypothesis testing (Gregory, 1990). Based on the available sensory information combined with prior experience of similar situations in the past, the most likely perceptual interpretation is selected. In the case of the rubber hand illusion, ownership of the fake hand is deemed more likely by the perceptual system than the alternative, which is that two perfectly synchronized sources of sensory stimulation are in fact unrelated. Other illusory percepts, such as the Ames room illusion, have also been taken as evidence for Gregory's theory. Here, based on prior experience and the available sensory evidence, a dramatic change of people's body size is deemed more likely than a room wall configuration which is trapezoidal rather than parallel.

In the rubber-hand illusion, people are sitting passively behind a table, and movement of the arm or hand is not allowed. In fact, if people nevertheless move their arm or hand, then the illusion will diminish or break. Yet, motor action and corresponding efferent and afferent information are equally important in establishing our bodily boundaries as is shown in several studies (for an overview, see, e.g., Jeannerod, 2003; Knoblich, 2002).

The capability of the central nervous system to incorporate tools and technological artifacts, including the components of immersive media technology, as a phenomenological extension of the self allows for exciting propositions. Consider the following excerpt from an interview with Lanier and Biocca (1992): "What if you took all the measurements and the movements of your physical body and somehow put them through a mathematical function that allowed you to learn to control six arms at once with practice? These sorts of things that play games with the feedback loop ... will be the real cutting edge of exploration of virtual reality as opposed to any particular symbolic content" (p. 162). Although the limits of what Lanier calls *homuncular flexibility* are still unknown, its promise is the experience of multiple atypical morphologies (see also Blakeslee and Blakeslee, 2007; Murray and Sixsmith, 1999). Perhaps in the future, it may even be possible to answer Nagel's (1974) question of what it is like to be a bat by ways of firsthand experience.

3. Discussion

In the present paper, we have outlined the first steps toward a theoretical framework that aims to explain the experience of telepresence from an embodied perspective. Our framework extends earlier embodied explanations of telepresence (e.g., Biocca, 1997; Ijsselstein, 2005; Slater and Usoh, 1994) by describing in detail how human beings are embodied. Based on Metzinger (2003, 2006) and Gallagher (1986, 2005a,b), our framework distinguishes three different levels of human embodiment: Morphology, body schema, and body image. The body schema is defined as a dynamic distributed network of procedures aimed at guiding behavior. The experience of being there in a virtual or mediated environment is explained through the ability of the network to adapt itself to the use of tools and technological artifacts. By doing so, the components of immersive media technologies, such as a head-mounted display, become integrated in a functional unity with our biological limbs and sensory receptors. As a result, we can interact with and through these technologies, as if they were not there. This is central to the fluent, intuitive interactions between humans and their technological surroundings, and forms the basis for interface transparency (Winograd and Flores, 1986) or cognitive disappearance (Weiser, 1991). The body image, in turn, is defined as being conscious of the body as owned by an individual, and includes not only our immediate bodily experience, but our emotions and memories that pertain to the body as well. Through incorporation in the body image, tools and technological artifacts can be experienced as an actual part of one's own body. To describe the profound effects that immersive media technology can have on our corporeal awareness and self-identity, Biocca (1997) coined the term *self-presence* (for a refinement of the term, see (Lee, 2004)).

In the present framework, the experience of telepresence depends largely on sensorimotor integration and establishing existing or new sensorimotor contingencies. These contingencies, to a large extent, are dependent on the way in which we are embodied. Media technologies that are interactive and rich (i.e., in terms of the level of detail and the number of sensory channels that are addressed) are more easily incorporated into our embodiment because they allow for more and more complex sensorimotor interactions. At the same time, our framework is consistent with the view that the experience of telepresence is an important measure of the quality of immersive media technology. A vivid experience of telepresence indicates that the media technology has been incorporated in the body schema, which in turn is necessary for a fluent interaction with and through those technologies.

Our framework makes a distinction between two types of incorporations: Functional extensions of the body (through incorporation in the body schema) and phenomenological extensions of the self (through incorporation in the body image). The benefit of a possible double incorporation is perhaps most obvious in the case of mechanical prostheses, where an amputee's attitude toward his or her prosthetic limb depends on both the functionality of the prosthesis (i.e., how well it can be operated), and the extent to which the prosthesis is experienced as a part of the self (Desmond and MacLachlan, 2002). Although both types of incorporations are expected to depend highly on sensorimotor integration, the exact underlying mechanisms may be different. The current evidence, for example, suggests that incorporation in the body schema requires active purposeful use of a tool (e.g., Yamamoto and Kitazawa, 2001; Iriki et al., 1996; Berti and Frassinetti, 2000; Cardinali et al., 2009). No active use of the external object is required for body image incorporations as is evident from the rubber-hand illusion, which may be evoked through visuotactile (Botvinick and Cohen, 1998) but also tactile-kinesthetic correla-

tions (Ehrsson et al., 2005). In contrast, body schema incorporations seem to require correlations between afferent information on the one hand, and information from the kinesthetic system and/or efference copy on the other. The question, then, remains why we experience ownership over some tools (such as a slave robot's arm) but not over others (such as a hammer). A possible answer is again provided by research on the rubber-hand illusion which demonstrates that the extent to which a foreign object is incorporated in the body image does not depend exclusively on a Bayesian inference of correlated multisensory inputs but also on the morphological congruence between the object and the human body (Tsakiris and Haggard, 2005; Haans et al., 2008). In other words, where the body schema is promiscuous, the body image appears to discriminate against tools that are morphologically different from the human body. Relevant in this respect is a recent experiment by Pusch et al. (2011) on the effects of different virtual hand representations on a pointing task. They found that a highly abstract pointer is to a large extent as effective functionally as a highly detailed and anthropomorphically correct hand model, although users generally experience a higher degree of control and comfort when using the latter.

The present framework summarizes and explains many of the key factors and processes behind telepresence, including interactivity, sensory-motor integration, media transparency, and distal attribution. As any theoretical account, however, it is not all encompassing. We have, for example, not focused on the role of attention and emotions in the experience of telepresence (e.g., Riva et al., 2007; Gorini et al., 2011). Second, the framework is applicable to the experience of presence in mediated and simulated environments. Whether it has any potential in explaining presence in imaginary environments (e.g., when dreaming, or reading a book) remains speculative (Biocca, 1997).

According to Loomis (1993), any satisfactory theory of telepresence should be grounded in the understanding of ordinary sensorimotor processes. We believe that our theoretical framework is a step towards such an understanding of telepresence. It is only because our body schema and self models function so flawlessly, that we often forget that all our actions and perceptions are mediated by them even in the absence of technology (cf. Loomis, 1992; Biocca, 1997). Tools and technologies are but additional mediators, which when appropriately integrated in our embodiment yield the same transparency that we experience when using our own natural sensors and actuators or, in a word, our bodies.

References

- Armell, K.C., Ramachandran, V.S., 2003. Projecting sensations to external objects: evidence from skin conductance response. *Proc. Roy. Soc. Lond. B Biol.* 270, 1499–1506.
- Berlucchi, G., Aglioti, S., 1997. The body in the brain: neural bases of corporeal awareness. *Trends Neurosci.* 20, 560–564.
- Berti, A., Frassinetti, F., 2000. When far becomes near: re-mapping of space by tool use. *J. Cogn. Neurosci.* 12, 415–420.
- Biocca, F., 1997. The cyborg's dilemma: progressive embodiment in virtual environments. *J. Comput-Mediat. Commun.* 3 (2).
- Biocca, F., Rolland, J., 1998. Virtual eyes can rearrange your body: adaptation to visual displacement in see-through, head-mounted displays. *Presence-Teleop. Virt.* 7, 262–277.
- Blakeslee, S., Blakeslee, M., 2007. *The Body Has a Mind of Its Own: How Body Maps in Your Brain Help You Do (Almost) Everything Better*. Random House, New York.
- Botvinick, M., 2004. Probing the neural basis of body ownership. *Science* 305, 782–783.
- Botvinick, M., Cohen, J., 1998. Rubber hands 'feel' touch that eyes see. *Nature* 391, 756.
- Brooks, R.A., 1991. Intelligence without representation. *Artif. Intell.* 47, 139–159.
- Cardinali, L., Frassinetti, F., Brozzoli, C., Urquizar, C., Roy, A.C., Farnè, A., 2009. Tool-use induces morphological updating of the body schema. *Curr. Biol.* 19, 478–479.
- Cole, J., Sacks, O., Waterman, I., 2000. On the immunity principle: a view from a robot. *Trends Cogn. Sci.* 4, 167.
- Dennett, D.C., 1991. *Consciousness Explained*. Little, Brown, Boston.
- Desmond, D., MacLachlan, M., 2002. Psychological issues in prosthetic and orthotic practice: a twenty-five year review of psychology in prosthetics & orthotics international. *Prosthet. Orthot. Int.* 26, 182–188.
- Dinh, H.Q., Walker, N., Song, C., Kobayashi, A., Hodges, L.F., 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In: Rosenblum, L., Astheimer, P., Teichmann, D. (Eds.), *P. IEEE Virt. Real.* IEEE Press, Los Alamos, CA, pp. 222–228.
- Dourish, P., 2001. *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press, Cambridge, UK.
- Edelman, G.M., 2003. Naturalizing consciousness: a theoretical framework. *Proc. Natl. Acad. Sci. USA* 100, 5520–5524.
- Edelman, G.M., 2006. *Second Nature: Brain Science and Human Knowledge*. Yale University Press, New Haven, CT.
- Ehrsson, H.H., Holmes, N.P., Passingham, R.E., 2005. Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *J. Neurosci.* 25, 10564–10573.
- Gallagher, S., 1986. Body image and body schema: a conceptual clarification. *J. Mind Behav.* 7, 541–554.
- Gallagher, S., 2005a. Dynamic models of body schematic processes. In: de Preester, H., Knockaert, V. (Eds.), *Body Image and Body Schema*. John Benjamins, Amsterdam, pp. 233–250.
- Gallagher, S., 2005b. *How the Body Shapes the Mind*. Oxford University Press, New York.
- Gallagher, S., Cole, J., 1995. Body schema and body image in a deafferented subject. *J. Mind Behav.* 16, 369–390.
- Gallup Jr., G.G., 1970. Chimpanzees: self-recognition. *Science* 167, 86–87.
- Gorini, A., Capideville, C.S., De Leo, G., Mantovani, F., Riva, G., 2011. The role of immersion and narrative in mediated presence: the virtual hospital experience. *Cyberpsychol. Behav. Soc. Netw.* 14, 99–105.
- Gregory, R.L., 1990. *Eye and Brain: The Psychology of Seeing*, fourth ed. Princeton University Press, Princeton, NJ.
- Groen, J., Werkhoven, P.J., 1998. Visuomotor adaptation to virtual hand position in interactive virtual environments. *Presence-Teleop. Virt.* 7, 429–446.
- Haans, A., IJsselstein, W.A., 2007. Self-attribution and telepresence. In: Moreno, L., Starlab Barcelona, S.L. (Eds.), *P. Ann. Int. Workshop Presence*, Starlab Barcelona S.L., Barcelona, Spain, pp. 51–58.
- Haans, A., IJsselstein, W.A., de Kort, Y.A.W., 2008. The effect of similarities in skin texture and hand shape on perceived ownership of a fake limb. *Body Image* 5, 389–394.
- Head, H., Holmes, G.M., 1911. Sensory disturbances from cerebral lesions. *Brain* 34, 102–254.
- Held, R., Durlach, N., 1991. Telepresence, time delay and adaptation. In: Ellis, S.R. (Ed.), *Pictorial Communication in Virtual and Real Environments*, second ed. Taylor & Francis, London, pp. 233–246.
- Held, R., Hein, A., 1958. Adaptation to disarranged hand-eye coordination contingent upon reafferent stimulation. *Percept. Motor Skills* 8, 87–90.
- Holmes, N.P., Spence, C., 2004. The body schema and multisensory representation(s) of peripersonal space. *Cogn. Process.* 5, 94–105.
- IJsselstein, W.A., 2005. Towards a neuropsychological basis of presence. *Annu. Rev. CyberTherapy Telemed.* 3, 25–30.
- IJsselstein, W.A., de Ridder, H., Freeman, J., Avons, S.E., 2000. Presence: concept, determinants and measurement. *Proc. SPIE* 3959, 520–529.
- IJsselstein, W.A., 2004. *Presence in Depth*. Eindhoven University of Technology, Eindhoven, The Netherlands.
- IJsselstein, W.A., de Kort, Y.A.W., Haans, A., 2006. Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence-Teleop. Virt.* 15, 455–464.
- Iriki, A., Tanaka, M., Iwamura, Y., 1996. Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 7, 2325–2330.
- Jeannerod, M., 2003. The mechanisms of self-recognition in humans. *Behav. Brain Res.* 142, 1–15.
- Knoblich, G., 2002. Self-recognition: body and action. *Trends Cogn. Sci.* 6, 447–449.
- Kugel, J., 1969. Lichaamsplan, lichaamsbesef, lichaamsidee: De psychologische betekenis van de lichamelijke ontwikkeling [Body Plan, Body Consciousness, and Body Idea: The Psychological Significance of Physical Development]. Wolters-Noordhoff, Groningen, The Netherlands.
- Lackner, J.R., 1988. Some proprioceptive influences on the perceptual representation of body shape and orientation. *Brain* 111, 281–297.
- Lanier, J., Biocca, F., 1992. An insider's view of the future of virtual reality. *J. Commun.* 42 (4), 150–172.
- Lee, K.M., 2004. Presence, explicated. *Commun. Theory* 14, 27–50.
- Lombard, M., Ditton, T., 1997. At the heart of it all: the concept of presence. *J. Comput-Mediat. Commun.* 3 (2).
- Loomis, J.M., 1992. Distal attribution and presence. *Presence-Teleop. Virt.* 1, 113–119.
- Loomis, J.M., 1993. Understanding synthetic experience must begin with the analysis of ordinary perceptual experience. In: *P. IEEE Symp. Res. Frontiers Virt.* IEEE Press, Los Alamos, CA, pp. 54–57.
- Maravita, A., Iriki, A., 2004. Tools for the body (schema). *Trends Cogn. Sci.* 8, 79–86.
- Maravita, A., Spence, C., Driver, J., 2003. Multisensory integration and the body schema: close to hand and within reach. *Curr. Biol.* 13, 531–539.
- Meltzoff, A.N., Moore, M.K., 1997. Explaining facial imitation: a theoretical model. *Early Dev. Parent.* 6, 179–192.

- Metzinger, T., 2003. *Being No One: The Self-model Theory of Subjectivity*. MIT Press, Cambridge, MA.
- Metzinger, T., 2006. Reply to Gallagher: different conceptions of embodiment. *Psyche* 12 (4).
- Metzinger, T., 2009. *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. Basic Books, New York.
- Minsky, M., 1980. Telepresence. *Omni* 21, 45–51.
- Minsky, M., 1988. *The Society of Mind*. Simon and Schuster, New York.
- Murray, C.D., Sixsmith, J., 1999. The corporeal body in virtual reality. *Ethos* 27, 315–343.
- Nagel, T., 1974. What is it like to be a bat? *Philos. Rev.* 83, 435–450.
- O'Regan, J.K., Myin, E., Noë, A., 2005. Skill, corporality and alerting capacity in an account of sensory consciousness. In: Laureys, S. (Ed.), *Progress in Brain Research, The Boundaries of Consciousness: Neurobiology and Neuropathology*, vol. 150. Elsevier, Amsterdam, The Netherlands, pp. 55–68.
- Plotnik, J.M., de Waal, F.B., Reiss, D., 2006. Self-recognition in an Asian elephant. *Proc. Natl. Acad. Sci. USA* 103, 17053–17057.
- Povinelli, D.J., Cant, J.G.H., 1995. Arboreal clambering and the evolution of self-conception. *Quart. Rev. Biol.* 70, 393–421.
- Prior, H., Schwarz, A., Güntürkün, O., 2008. Mirror-induced behavior in the Magpie (*Pica pica*): evidence of self-recognition. *PLoS Biol.* 6, 1642–1650.
- Pusch, A., Martin, O., Coquillart, S., 2011. Effects of hand feedback fidelity on near space pointing performance and user acceptance. In: Park, J.-I., Wan, W., Kajimoto, H., Nii, H. (Eds.), *P. IEEE Inter. Symp. Virt. Real. Innov. IEEE Press, Los Alamos, CA*, pp. 97–102.
- Reiss, D., Marino, L., 2001. Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence. *Proc. Natl. Acad. Sci. USA* 98, 5937–5942.
- Riva, G., Mantovani, F., Capideville, C.M., Preziosa, A., Morganti, F., Villani, D., Gaggioli, A., Botella, C., Alcaniz, M., 2007. Affective interactions using virtual reality: the link between presence and emotions. *Cyberpsychol. Behav.* 10, 45–56.
- Rochat, P., Striano, T., 2000. Perceived self in infancy. *Infant Behav. Dev.* 23, 513–530.
- Sheridan, T.B., 1992. Musings on telepresence and virtual presence. *Presence-Teleop. Virt.* 1, 120–125.
- Shore, D.I., Spry, E., Spence, C., 2002. Confusing the mind by crossing the hands. *Cogn. Brain Res.* 14, 153–163.
- Slater, M., 2003. A note on presence terminology. *Presence-Connect* 3 (January).
- Slater, M., 2004. Presence and emotions. *Cyberpsychol. Behav.* 7, 45–56.
- Slater, M., Usoh, M., 1994. Body centred interaction in immersive virtual environments. In: Magnenat Thalmann, N., Thalmann, D. (Eds.), *Artificial Life and Virtual Reality*. John Wiley, Chichester, UK, pp. 125–148.
- Slater, M., Perez-Marcos, D., Ehrsson, H., Sanchez-Vives, M.V., 2008. Towards a digital body: the virtual arm illusion. *Front. Hum. Neurosci.* 2 (6).
- Stanney, K.M., Kennedy, R.S., Drexler, J.M., Harm, D.L., 1999. Motion sickness and proprioceptive aftereffects following virtual environment exposure. *Appl. Ergon.* 30, 27–38.
- Steuer, J., 1992. Defining virtual reality: dimensions determining telepresence. *J. Commun.* 42 (4), 73–93.
- Strauss, S., 1995. Virtual Reality Too Real for many. *Toronto Globe & Mail*, pp. A1, A8 (March 4).
- Tsakiris, M., Haggard, P., 2005. The rubber hand illusion revisited: visuotactile integration and self-attribution. *J. Exp. Psychol. Human* 31, 80–91.
- Varela, F.J., Thompson, E., Rosch, E., 1991. *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press, Cambridge, UK.
- Weiser, W., 1991. The computer for the twenty-first century. *Sci. Am.* 265 (3), 94–104.
- Welch, R.B., 1978. *Perceptual Modification: Adapting to Altered Sensory Environments*. Academic Press, New York.
- Welch, R., Blackmon, T., Liu, A., Mellers, B., Stark, L., 1996. The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence-Teleop. Virt.* 5, 263–273.
- Winograd, T., Flores, F., 1986. *Understanding Computers and Cognition: A New Foundation for Design*. Ablex, Norwood, NJ.
- Yamamoto, S., Kitazawa, S., 2001. Reversal of subjective temporal order due to arm crossing. *Nat. Neurosci.* 4, 759–765.
- Yamamoto, S., Kitazawa, S., 2001. Sensation at the tips of invisible tools. *Nat. Neurosci.* 4, 979–980.
- Zahorik, P., Jenison, R.L., 1998. Presence as being-in-the-world. *Presence-Teleop. Virt.* 7, 78–89.