User Representations in Human-Computer Interaction

Sofia Seinfeld, Tiare Feuchtner, Antonella Maselli & Jörg Müller

Taylor & Francis
Taylor & Francis Group

 OPEN ACCESS   Check for updates

User Representations in Human-Computer Interaction

Sofia Seinfeld ⓘ[a], Tiare Feuchtner[b], Antonella Maselli[c], and Jörg Müller[a]

[a]Institute of Computer Science, University of Bayreuth, Bayreuth, Germany; [b]Department of Computer Science, Aarhus University, Aarhus, Denmark; [c]Department of Neuromotor Physiology, Santa Lucia Foundation, Rome, Italy

ABSTRACT

Cursors, avatars, virtual hands or tools, and other rendered graphical objects, enable users to interact with computers such as PCs, game consoles or virtual reality systems. We analyze the role of these various objects from a user perspective under the unifying concept of "User Representations". These representations are virtual objects that artificially extend the users' physical bodies, enabling them to manipulate the virtual environment by performing motor actions that are continuously mapped to their User Representations. In this paper, we identify a set of concepts that are relevant for different User Representations, and provide a multidisciplinary review of the multisensory and cognitive factors underlying the control and subjective experience of User Representations. These concepts include visual appearance, multimodal feedback, sense of agency, input methods, peripersonal space, visual perspective, and body ownership. We further suggest a research agenda for these concepts, which can lead the human-computer interaction community towards a wider perspective of how users perceive and interact through their User Representations.

1. Introduction

We are currently witnessing an unprecedented diversification of computer interfaces. Nowadays, users are able to interact with virtual worlds through diverse means beyond the usual mouse cursors. For example, when playing video games, users are often provided with a virtual avatar experienced from either a first person (1PP) or third person perspective (3PP). The avatar's visual style may vary from a photorealistic appearance to a cartoonish one. With the development of augmented reality (AR) and virtual reality (VR) interfaces, it has further become possible to manipulate virtual objects via realistic collocated representations of virtual controllers, tools, hands or even full virtual avatars.

Historically, no overarching analysis of these different means of interaction has been conducted, and we lack a deeper understanding of the cognitive basis and behavioral correlates of differently representing the user in a given interface. To address this issue, we propose the unifying concept of "User Representations" and describe diverse means of interaction as instantiations of this common concept. With User Representations we refer to virtual objects that extend the users' physical bodies into virtual environments, enabling them to execute actions there. Importantly, users are able to continuously control these virtual User Representations through their motor commands.

This integrated perspective allows us to discuss diverse virtual representations and interfaces within a common framework, and to apply concepts from cognitive neuroscience for analysis and classification. Such concepts include visual appearance, multimodal feedback, sense of agency, input method, peripersonal space, visual perspective, and body ownership.

In direct manipulation involving continuous representation of the objects of interest, physical actions, and rapid incremental reversible operations (Hutchins, Hollan, & Norman, 1985;
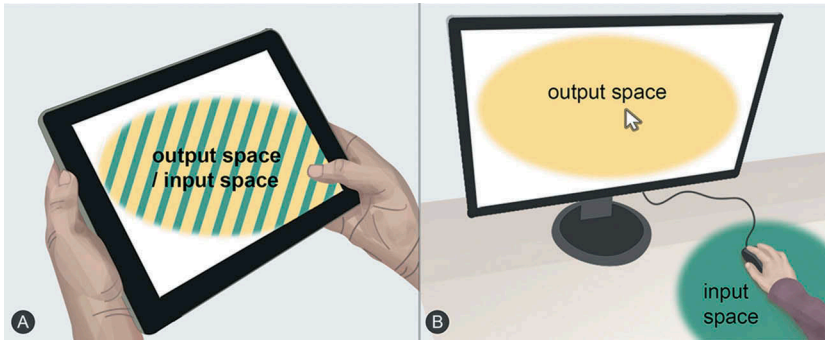

CONTACT Sofia Seinfeld   ✉ sofia.seinfeld@uni-bayreuth.de   Institute of Computer Science, University of Bayreuth, Bayreuth, Germany

Figure 1. Interaction with Graphical User Interfaces can either be direct or indirect, depending on the mapping between input and output space. In touchscreen interaction, the input and output space are collocated on the screen itself ((A): Direct Interaction), while mouse interaction involves a separation of input space (mouse on a desk) and output space (cursor on computer screen) ((B): Indirect Interaction).

Shneiderman, 1997), User Representations form a crucial element for interaction with virtual environments. This is particularly true for *indirect interactions*, where input and output space are separated. In this context, the input space refers to the part or area of an interface where physical actions may be carried out to interact in a given virtual environment. The output space is the location where virtual objects are visualized and manipulated (e.g., a display or a screen). For example, when users activate a mouse to move a cursor on a screen in order to select a virtual object, the input space corresponds to the area on which the mouse is handled, while the output space is the computer screen where the virtual objects are visualized. However, not all user interfaces have this separation of input and output space. For instance, when users select a virtual object on a touchscreen (e.g., tablet or smartphone) directly through contact with a fingertip (Figure 1A), the input and output spaces are collocated. We refer to this phenomenon as *direct interaction*. In contrast, during indirect interactions users are not able to directly select virtual objects with their fingertips, but require a virtual representation (e.g., a cursor) to manipulate virtual objects. Thus, the cursor is the User Representation that enables actions with other virtual objects, in *indirect interaction* (Figure 1B).
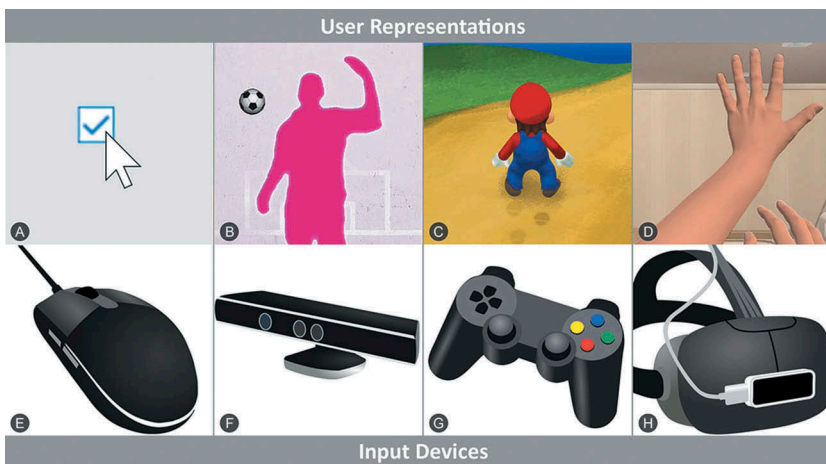

Figure 2. User Representations vary in their appearance and the typical input devices and mapping functions that are used to control them. While many different combinations of these characteristics are possible, some combinations are more common, such as mouse and pointer (A-E), Kinect and mirror image (B-F), gamepad and 3D avatar (C-G), or Leap Motion and virtual hands (D-H).

In this paper we focus solely on indirect interaction, where a virtual object, namely a User Representation, indicates where and how the user's manipulations will take effect in the virtual world.

Examples of User Representations include: i) an arrow-shaped mouse cursor in a desktop environment (Figure 2A), ii) an avatar seen from a 1PP or 3PP on a screen (Denisova & Cairns, 2015) (Figure 2C) or in immersive VR (Galvan Debarba et al., 2017) (Figure 2D), iii) the user's mirror image in a virtual world, for example when interacting with a large display (Gonzalez-Franco, Perez-Marcos, Spanlang, & Slater, 2010; Müller, Walter, Bailly, Nischt, & Alt, 2012; Slater, Spanlang, Sanchez-Vives, & Blanke, 2010) (Figure 2B). We can also envision novel and future methods of representing the user, such as dynamic physical interfaces where users can be represented through physically levitating particles that they control in real time (Bachynskyi, Paneva, & Müller, 2018; Hirayama, Martinez Plasencia, Masuda, & Subramanian, 2019), humanoid robots (Kishore, Navarro, Dominguez, De La Peña, & Slater, 2018), or drones (Fernandez et al., 2016).

As can be appreciated in the abovementioned examples, User Representations may differ in various characteristics. For example, they may vary in their morphological characteristics, the input device used to control them, the mapping established to transfer motor actions into state changes of the representation, the experienced visual perspective (i.e., 1PP or 3PP), the type of multimodal feedback included, whether the representation is clearly experienced as an external tool or as part of the body (i.e. illusory body ownership). Despite these differences, a common principle underlying interaction through User Representations is that users perceive these virtual objects as representations of themselves and feel that they can effectively perform actions through them.

In the last few years, a vast amount of research in psychology, neuroscience, and Human Computer Interaction (HCI) has yielded knowledge on factors that can impact the perception and control of User Representations. Relevant topics such as the embodiment of manipulated tools, peripersonal space, sense of agency, affordances, sense of ownership, and multisensory integration, have been investigated extensively. However, the vast majority of these studies have not addressed the role of these factors with reference to User Representations. Moreover, there has been no critical analysis of how variations in each of these factors could impact on interaction via a virtual object that represents the user in a given interface. The present paper aims to remedy this situation, by providing an integrated framework for User Representations based on these factors. Such an organized and integrated perspective can yield deeper insights into how human cognitive schemas and skills are impacted by different User Representations, and how this may be exploited to improve human interactions through their virtual counterparts. For example, an understanding of this framework may allow researchers to critically evaluate how the visualization and control-mapping function for a User Representation can be adapted to better support a specific task, how user perceptions can be tricked into perceiving different physical properties (e.g., stiffness, weight) when manipulating virtual objects through a User Representation, or how the embodiment of a User Representation can be leveraged to promote attitudinal and behavioral changes.

Thus, the main objectives of the present article are: i) to define User Representations based on their common underlying principles (Section 2); ii) to review relevant research and concepts from HCI, psychology, and neuroscience, which impact on the perception and control of User Representations (Section 3); iii) to illustrate with examples, how these concepts serve to analyze classic and new interfaces (Section 4); iv) to provide a framework for User Representations based on the reviewed concepts and to establish a research agenda that advances future knowledge on interaction and User Representations (Section 5).

2. Definition and characterization of user representations

2.1. *Definition*

The present article defines User Representations in terms of two basic principles which are common to all types of User Representations, as follows:

(1) These virtual objects or User Representations act as artificial extensions of users' physical bodies, enabling them to execute actions in virtual environments, which otherwise would be unreachable.

For example, in a traditional desktop environment, users select and deselect different virtual objects viewed on a screen, through the use of a cursor, but not using their hands directly to touch and grab these virtual objects. When playing a desktop videogame, users are frequently represented by an avatar that carries out different actions in the virtual environment in order to play the game.

(2) Users are able to continuously control these virtual objects or User Representations through their motor commands, resulting in a perceived sense of agency over the representations.

These motor commands include hand movements, button pressing, and full-body motions. Such sensorimotor contingencies (i.e., the spatiotemporal correlations among sensory cues and motor commands or actions), result in users perceiving a sense of agency over the actions executed through the User Representation. Accordingly, the users perceive themselves as the agents responsible for the actions carried out by the User Representation they are controlling.

2.2. Characterizing user representations from HCI interaction models

Interaction between humans and computers has been described from various perspectives. These include models that understand interaction as tool use (Bødker, 1987; Klokmose & Beaudouin-Lafon, 2009), as control (Müller, Oulasvirta, & Murray-Smith, 2017), or as a dialogue (Card, Mackinlay, & Robertson, 1990), among others. These perspectives are relevant for understanding the proposed concept of User Representations.

2.2.1. From a tool-use perspective

From a tool-use perspective, User Representations can be viewed as tools that a user manipulates in order to generate changes in a given interface. Thus, returning to the above definition, they act as an artificial extension of the user's physical body. This is analogous to the use of any other type of tool for accomplishing a task (e.g., hammer, screwdriver, stick, and shovel). When considering interaction as tool-use, three major implications have been highlighted (for a detailed explanation, see Hornbæk & Oulasvirta, 2017). The first implication is that tools shape user perceptions and behavior, as also suggested by activity theory and instrumental interaction. The present article aims to expand this view from the standpoint of psychology and cognitive neuroscience. We discuss how user behavior and cognition are differently influenced by virtual-tool use depending on several characteristics of the User Representation (e.g., their visual features, spatial configurations, the provided sensorimotor contingencies between input device manipulations and resulting state changes of the representation, the inclusion of additional multimodal feedback, the anthropomorphic characteristics of the representation, among others).

The second implication underscores that tools play a mediating role. Here we consider that User Representations, controlled through an input device, play a fundamental role in enabling the user to interact with the interface. For example, Alzayat, Hancock, and Nacenta (2019) highlight the particular importance of tools in HCI, and provide a VR example. VR interactions are frequently mediated by one or multiple input devices such as the HTC Vive Controllers. In several VR applications we can see the virtual counterpart of these controllers, so that they are also represented as tools in the virtual world. These VR tools mediate user interactions with other virtual objects, for example when the person selects other virtual objects or paints, using the virtual version of the controller (see VR painting applications in Section 4). The third implication is related to the usefulness and directness of the tool. Our framework discusses these characteristics with reference to different types of User Representations.
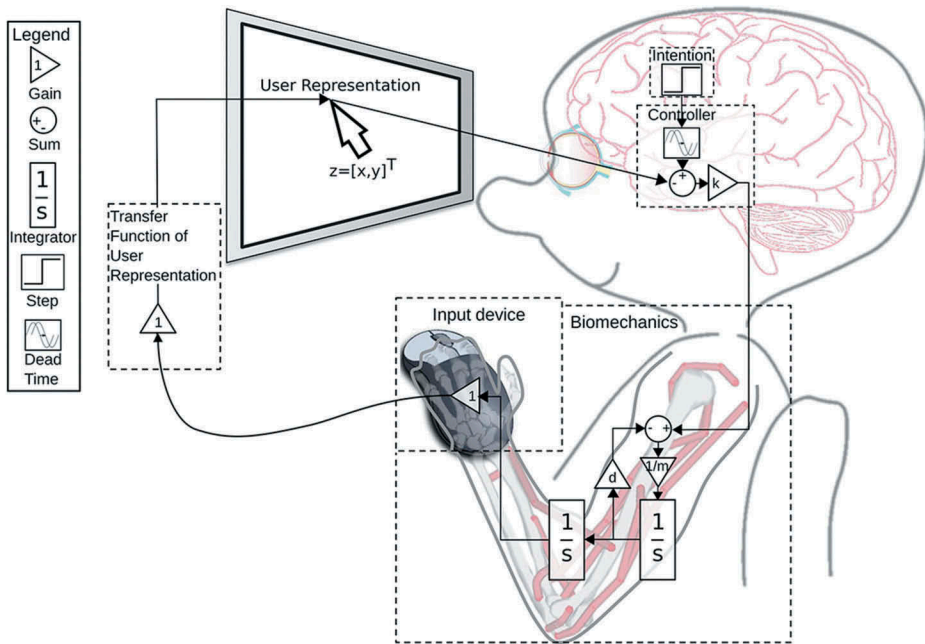
Figure 3. The User Representation can be interpreted as a dynamical system of which users can control the state, such as the position of the mouse cursor on the screen. Users control the state through body movements, which are measured by input devices. Users continuously observe the state of the User Representation and adjust their movement in order to achieve some desired state, such as moving the User Representation towards a target area. In this way, a continuous control loop between the user and the User Representation emerges.

Finally, considering User Representations as virtual tools raises the important question of how these differ from other types of virtual tools included in an interface (e.g., icons, scrollbars, etc.). From the perspective of the present framework the main difference is that virtual tools provide specific functions to manipulate objects of interest in individual applications. By contrast, a User Representation is controlled continuously by the user during interaction, and can provide multiple functionalities across multiple applications. Such continuous control is achieved through specific motor actions (e.g., hand movements) and enables users to know where and when their actions take place in the interface. A further distinction is that an interface may contain several virtual tools, whereas the user is typically represented by only one User Representation at a time.

2.2.2. *From a control perspective*

The dynamical systems perspective of interaction (Müller et al., 2017) models the relationship between human and computer as a control loop, as illustrated in Figure 3. From this perspective, the User Representation can be seen as a dynamical system. The user observes and controls the state of this system through control signals mediated by an input device. The state of the User Representation can be described by a vector z of state variables, which change continuously over time. The state of a mouse pointer, for example, simply encodes the current position of the pointer as $z = [x,y]T$, for which the state variables x and y are the screen coordinates of the pointer along the x and y axes. Similarly, the state of a virtual avatar might consist of the position and orientation of the virtual body, and all of the joint angles.

2.2.3. *From a dialogue perspective*

The dialogue perspective of interaction (Card et al., 1990; Hutchins et al., 1985) views the interaction between humans and computers as a dialogue between them. In other words, the user needs to communicate with the computer application in order to get some tasks performed. Users must express their intentions in the

language of the computer interface. With regard to this perspective, the User Representation can be interpreted as the part of the interface that is used to define at which virtual object the dialogue is directed, so that it acts as the media structure which mediates communication between the human and the computer.

3. Psychology and neuroscience background: key cognitive concepts in relation to User Representations

3.1. *Visual appearance and affordances*

The concept of affordance has found widespread acceptance in HCI and has strongly influenced interface design (Bærentsen & Trettvik, 2002; Kaptelinin & Nardi, 2012). The term was originally coined by Gibson (1966) to describe the perceptual relation between the perceiver and the environment. From the ecological perspective, an affordance highlights the information that perceivers pick up from the environment in order to specify the potential actions of different objects and places (Gibson, 1977, 1979). For example, a small gap in a wall may enable a child to walk through, but not an adult. Norman (1988) introduced the concept to designers and programmers as the properties that allow users to determine how something could potentially be used. From his perspective, affordances can be divided into real (physical) or perceived (Norman, 1999). Norman (2008) further describes that what we perceive when we notice an affordance is a "signifier", which entails any other information that provides clues for making sense of something (e.g., object, design, situation).

The present article proposes that concepts such as affordances and signifiers can also be applied to User Representations in order to indicate the range of possible actions that the user can perform within a given interface or virtual world. In the context of this paper, we regard affordances as the possibilities for action between user and virtual environment, which are mediated by the User Representation. At the same time, the User Representation and other objects in the virtual world may have signifiers that communicate the actions they support. An example is a virtual hole that enables access to a small avatar, but not to a larger one. Another interesting aspect is that the visual properties of a User Representation can also impact on how the virtual environment is perceived. A good example is how individuals use the relative size of virtual hands that represent them within the interface, so as to establish a scale for the apparent size of the remaining virtual objects in the virtual environment (Linkenauger, Leyrer, Bülthoff, & Mohler, 2013; Ogawa, Narumi, & Hirose, 2019). Thus, virtual objects are perceived as larger as a function of decreases in virtual hand size.

Additionally, the symbolic meaning linked to the visual appearance of a User Representation, may also provide the user with valuable indications of the functions and actions afforded by that representation. Depending on the application we are using, the visual shape of a User Representation can change dynamically and take on radically different appearances according to the diverse functional roles associated with them. The following list provides a broad classification of possible visual appearances (Kadri, Lecuyer, Burkhardt, & Richir, 2007):

- Virtual object without well-defined symbolic meaning or "signifiers" (e.g., arrow, sphere, line, triangle).
- Representation depicting a specific tool related to its function (e.g., scissors, eraser, paintbrush, hammer, etc.).
- Representation of a body part (e.g., eyes, hands, fingers).
- Full-body avatar representation, which can range from an extremely realistic avatar to a cartoon-like avatar.

Despite the scarcity of research exploring the relations between a User Representation's visual appearance and user behavior, those studies that have tackled this question have produced interesting results. For example, the visual features of an avatar or a three-dimensional (3D) cursor, influence the way users manipulate an input device to control their virtual representation (Kadri

et al., 2007). Moreover, it has also been shown that users change their hand position when controlling an input device, depending on the visual orientation of a virtual cursor. Similarly, in a menu-selection task, cursors including clear directional cues (e.g. arrows pointing in a certain direction) can differently impact user performance (Po, Fisher, & Booth, 2005). This shows that visual appearance and stimulus-response compatibility, influence the way virtual objects are manipulated in a virtual environment.

In fact, visual cues are exploited in direct manipulation of post-WIMP interfaces (Beaudouin-Lafon, 2000). Here, the functional role associated with User Representations frequently matches the attributes of the physical tools used to carry out analogous tasks in real life (Kato, Billinghurst, Poupyrev, Imamoto, & Tachibana, 2000). For instance, users can infer that a User Representation depicting scissors will enable them to cut certain content depicted on screen, while the selection of a virtual paintbrush will enable them to paint a virtual surface. Video games provide a further example of communicating the status of the game to the player by varying the appearance of the avatar. In Super Mario Land[1] the character increases in size after "eating" a mushroom, which signifies to the player that he will now have the power to break bricks with his head; when collecting a star, a blinking animation indicates the duration during which Mario is invulnerable. In newer games, avatars are frequently colored red or appear disheveled when their health is poor.

User Representations can differ radically in their visual appearance and functionalities. Depending on the application used, representations can take the shape of abstract objects, tools, body parts or a full-body avatar. These visual characteristics of the representation can also quickly and dynamically change, so as to signify different functionalities. For example, a cursor can change its appearance dynamically, depending on its location on a screen and the associated function linked to that location (e.g., a cursor taking the shape of a hand for dragging/moving, of an arrow for selection purposes, or of a caret to indicate that text can be entered). Similarly, a game character can change its visual features to signify its current state and skills. These visual properties have the potential to impact on subsequent user actions, since based on the appearance of the representation, the user may try to defeat a powerful villain allied with the superpower or win health points to recover from injuries.

3.2. *Multimodal feedback*

Human perception of the world is constructed in real time, based on the integration of different sensory modalities by the nervous system (Calvert, Spence, & Barry, 2004). Perception arises as a result of specialized mechanisms in the brain for integrating different sensory modalities such as smell, taste, vision, auditory, and somatosensory information (i.e. perception of touch, pressure, pain, and temperature). Additionally, information about body posture and movements (i.e., known as proprioception and conveyed from receptors in muscles and joints) is also integrated. Multisensory integration mechanisms are crucial for adaptive behavior, since they enable coherent and meaningful unitary perception of the world (Ernst & Bülthoff, 2004). When playing a tennis game, our control of the racket depends on the simultaneous integration of visual, auditory, proprioceptive, and motor information. We need to quickly modify our movements and body postures, based on the motion and sounds we perceive from the tennis ball, racket, and the other player. The combination of different modalities allows us to lower detection thresholds and improve spatial and temporal accuracy. In this regard, while vision plays a crucial role in estimating the position of the ball, the sound provides a more accurate estimation of the exact timing of the hit.

Besides enhancing the detection and perceptual accuracy of real events perceived through multiple sensory channels, multisensory integration is at the root of several perceptual illusions. Well known examples can be found in speech perception, where the visual aspects of speech cues have a significant impact on the corresponding heard speech. An example is the ventriloquist effect, where the location of

---

[1]https://nintendo.fandom.com/wiki/Super_Mario_Land.

a heard sound is misperceived toward the location of the visually perceived speech (Alais & Burr, 2004; Burr & Alais, 2006). In the McGurk effect instead, a sound perceived through lip reading, together with a different simultaneously heard sound, results in the perceptual illusion of a different third sound. Research has shown that these effects can be explained in terms of probabilistically optimal integration mechanisms following Bayes' Rule (Chen & Spence, 2017; Körding & Wolpert, 2004).

Similar effects can be exploited in HCI, for example to compensate for intrinsic temporal lags in the system, or when interacting through a User Representation. In fact, auditory, visual and haptic systems may have different latencies, intrinsically inherent to the associated rendering technology. We believe that a thoughtful consideration of the spatiotemporal and causal binding principles that govern multimodal perception may provide important guidelines to the HCI community, on how to adjust the interplay of different modality feedback, so to enhance the perceived naturalness of the interaction (Picard & Friston, 2014). However, HCI researchers may also consider exploiting subtle sensory misalignments, for example for pseudo-haptics or haptic retargeting (see Section 3.4 *Motor control, input methods, and mappings* for a detailed explanation). Recently researchers have also exploited extreme sensory misalignments in order to create thrilling experiences such as feeling we are virtually on a roller coaster (Marshall, Benford, Byrne, & Tennent, 2019).

In today's interfaces, the most common feedback provided by a User Representation is visual and auditory. However, multimodal user interfaces aim to take advantage of the richness of different sensory information. Senses like touch, smell, and even taste are stimulated during interactive tasks with technology (Obrist et al., 2016). As just one example, Akamatsu, Sato, and MacKenzie (1994) created a multimodal mouse that provides visuo-motor, tactile, and force feedback when a target is reached with the cursor, improving target-selection times, compared to a mouse that only offers visual feedback. In VR, researchers have also developed applications that provide haptic and force feedback when interacting with virtual objects, in addition to visual feedback (Georgiou et al., 2018; Salisbury & Srinivasan, 1997; Swapp, Pawar, & Loscos, 2006). The inclusion of this somatosensory feedback, which resembles the way we perceive the real world, has the potential to enhance user experiences and to help users identify with their artificial representations (Slater, 2009). However, it should be empathized that recent studies have also shown that sensory feedback has to be implemented very carefully in order to avoid undesired effects, such as the "uncanny valley" of haptics (Berger, Gonzalez-Franco, Ofek, & Hinckley, 2018).

3.3. *Sense of agency*

The sense of agency is the subjective experience of controlling the course of events through actions executed in the external world (Haggard, 2017). In the context of HCI, it refers to the extent we feel responsible for changing the state of a computer through our actions (Limerick, Coyle, & Moore, 2014). This is a key aspect of interface design, since users strongly desire to perceive that they are able to control technology as intended (Cornelio-Martinez, De Pirro, Vi, & Subramanian, 2017; Shneiderman, 1997). The sense of agency is modulated by several factors that can be exploited to enhance (or diminish) the sense of control users perceive over their virtual representations.

Substantial evidence supports the view that the feeling of agency over one's own actions is associated with the expected sensory feedback. Agency arises when there is a correspondence between the predicted sensory consequences of an action and the actual perceived changes in the environment produced by the action (Haggard, 2017; Limerick et al., 2014). This perspective has been successfully modeled by research focused on motor control. In the motor control community, several studies indicate that when the motor system formulates a plan for an action, it sends out a motor (execution) command to the body and simultaneously runs an internal model to simulate the planned action, also known as the efference copy (Figure 4). This internal model generates predictions for the associated sensory consequences of the action (Miall & Wolpert, 1996; Wolpert & Ghahramani, 2000). When the predicted and perceived consequences of an action match, a mutual
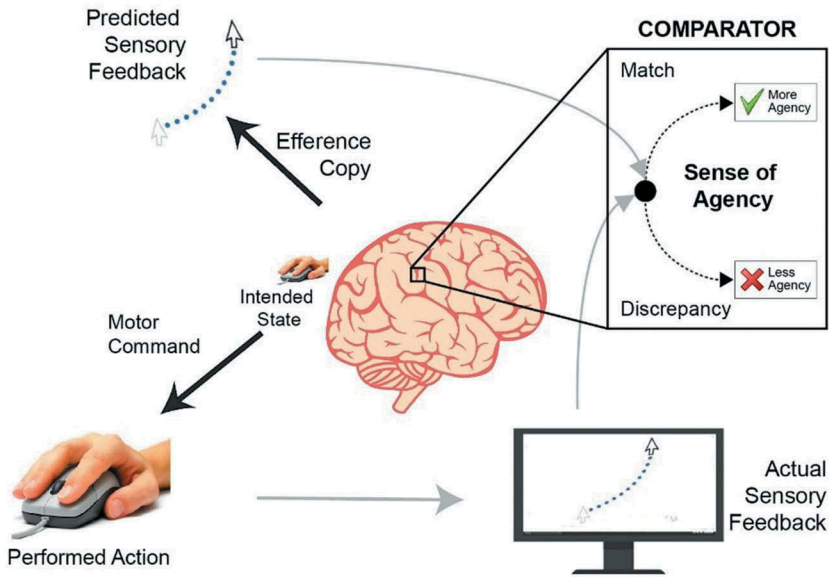
Figure 4. According to the Sense of Agency Model, when there is an intention for action (e.g., control of a cursor on a computer through a mouse), a motor command is issued to execute a movement. In parallel, in the brain an internal model simulates the action (efference copy). This model predicts sensory feedback expected from the executed action. If there is a match between predicted and perceived sensory feedback, then the sense of agency arises.

cause-effect relationship is established, giving rise to the sense of agency (Frith, Blakemore, & Wolpert, 2000; Picard & Friston, 2014).

The sensorimotor control perspective can explain the different forms of agency. In the case of body agency, it explains the sense of agency experienced for one's own body movements, as well as the feeling of controlling a virtual body. Visual exposure to virtual bodies or hands, animated in real time to follow the user's actual executed movements, is a powerful method for triggering feelings of agency over the seen virtual body (Slater et al., 2010). However, the sense of agency is not body-specific. This feeling of control is preserved when, instead of seeing a virtual hand, users see a virtual sphere (or other virtual objects) mimicking their own hand movements (Zopf, Polito, & Moore, 2018). Furthermore, contingent visuo-motor cues can also induce a strong feeling of agency, even when the mapping between motor commands and the experienced feedback are not spatiotemporally aligned, but still correspond to a known on-line mapping. This is the case, for example, when users see a cursor on a screen following their hand movements (Gozli & Brown, 2011). Here the user has a strong sense of being the agent responsible for moving the cursor and the events taking place in the virtual environment, even though hand and cursor are not collocated (Limerick et al., 2014).

Beside low-level sensory factors associated with the principles of sensorimotor control, other cognitive factors, such as goal-directed behaviors or prior knowledge, can have an important impact on the sense of agency (Limerick et al., 2014). For example, Kumar and Srinivasan (2014) found that when users play a multi-agent game, they still feel a high degree of control, even in conditions where they do not have full control of an agent, if the main goal of the game is accomplished by the other agents. This shows that when users are endowed with a high level of control (e.g., achieving a specific goal), the sense of agency seems to be less strongly modulated by slight variations in low-level sensory aspects.

Similar mechanisms may be activated in auto-aiming in games or pointing facilitation methods (Balakrishnan, 2004). With some of these methods, such as pointer acceleration (Casiez, Vogel, Balakrishnan, & Cockburn, 2008) or semantic pointing (Blanch, Guiard, & Beaudouin-Lafon, 2004), the cursor moves differently from the control exerted by the user, in order to improve pointing

performance. Coyle, Moore, Kristensson, Fletcher, and Blackwell (2012) investigated the impact of a machine-assisted point-and-click task on the user's sense of agency. Participants were instructed to reach targets with a mouse pointer in an application that included an algorithm to provide the user with different levels of assistance. The algorithm consisted of the addition of "gravity" to the targets, so as to attract the mouse pointer to the closest targets. It was found that when no or mild levels of assistance were provided, users experienced a high level of agency. However, a significant loss of agency was observed in medium and high levels of assistance conditions, although participants accomplished their goals (e.g., touching a target). Thus, it seems that the computer can assist users, but only up to a certain degree without negatively impacting agency over a User Representation. Very high levels of assistance seem to negatively influence users' sense of control.

However, agency is not only modulated by sensory and motor information, the cause of an event is also assessed given certain contextual cues. When environmental cues match the explanation that "I am" the agent of a given event, it is even possible to perceive agency over an action "I" did NOT execute (Synofzik, Vosgerau, & Newen, 2008). This sense of "vicarious" agency can occasionally occur when an unintended event, e.g., the switching on of a light, occurs at the same time of a willed independent movement (Michotte, 2017). For example, participants can experience agency over the arm movements of another person, when they see the other person's arms in a configuration that is congruent with the rest of their body and they simultaneously listen to audio-instructions that are consistent with respect to the seen movements (Wegner, Sparrow, & Winerman, 2004). Similarly, participants can also wrongly self-attribute the action of speaking, when they experience the illusion of being in the body of a speaking avatar, and their throat is synchronously stimulated (Banakou & Slater, 2014). Another study showed that is possible to experience the illusion of agency with respect to the walking of a virtual body, although participants are in a seated position and not executing any motor action. This is accomplished when participants have a 1PP view of a walking life-size virtual body seen through a head-mounted display (HMD) (Kokkinara, Kilteni, Blom, & Slater, 2016).

Agency can be objectively measured using a paradigm known as temporal binding. In HCI, this method has been adapted recently to assess users' feelings of agency during the interaction with different types of interfaces (Cornelio-Martinez, Maggioni, Hornbæk, Obrist, & Subramanian, 2018; Limerick et al., 2014; Moore & Obhi, 2012). Temporal binding is a phenomenon consisting of the underestimation of the time elapsed between an action and an outcome, only when the action is performed voluntarily and not involuntarily (Haggard, Clark, & Kalogeras, 2002). Using this measure, it has been found that there is a stronger temporal binding effect when an interface uses skin-based input (e.g., touching one's own arm) compared to pressing a button on a keyboard or touchpad (Bergstrom-Lehtovirta, Mottelson, Muresan, & Hornbæk, 2019; Coyle et al., 2012). It has also been shown that speech-based interfaces can lead to a diminished sense of agency or lower temporal binding effects, when compared to keyboard inputs (Limerick, Moore, & Coyle, 2015). These results suggest that the input method used to control User Representations will also impact on the degree of experienced agency during the interaction.

3.4. *Motor control, input methods, and mappings*

Users continuously control the state (e.g., position) of their User Representations through their body movements. For example, in the case of aimed movements, users aim to move some part of the User Representation, the end-effector, toward a spatially defined target as quickly as possible. Users continuously observe the state of the User Representation and move such that the difference between the current state of the User Representation and their goal state is minimized. Because their movement changes the state of the User Representation, and the changed state of the User Representation in turn changes the users' control movements, a closed control loop emerges. The users' movements are captured by input devices and passed on to the computer, for example to control the position of the mouse cursor on the screen. In this sense, input devices are "transducers from the physical properties of the world into logical values of an application" (Baecker & Buxton,

1987). Input devices have different sensors for measuring physical properties, such as the movement of a mouse on a surface, the pressure exerted on a button, or the tilt angle of a joystick. In some cases, the input device may contain sensors that can directly interpret the position and orientation of the users' body parts in space (e.g., Kinect,[2] Leap Motion,[3] Optitrack[4]). One (continuous) dimension that is sensed by the device and can be independently controlled by the user is often called a Degree of Freedom (DoF). For example, a mouse allows controlling 2 DoF by moving it along two dimensions in physical space (Figure 2E), leading to a corresponding movement of a mouse cursor (Figure 2A) in virtual space. Similarly, a joystick on a gamepad (Figure 2G) also permits the control of 2 DoF, but the gamepad can include multiple joysticks to provide control of more DoF, enabling complex movements of a virtual avatar (Figure 2C). Furthermore, motion tracking sensors like the Kinect and Leap Motion (Figure 2F,H) can measure joint angles of the user's body, thus allowing control of a large number of DoFs that reflect a significant fraction of human full-body biomechanics. Thereby, the user's natural body movements can be faithfully reflected by a virtual avatar (Figure 2B,D). The mapping between input (e.g., mouse movement) and output (cursor movement) is commonly called a transfer function (Casiez et al., 2008).

In some cases, the input device may provide control of fewer DoF than the User Representation offers. In such cases, multiple input devices can be combined, each providing some of the available DoF. For example, in a shooter game, the player may only be able to control the avatar's viewing direction with a mouse or joystick, and a second input device (e.g., use of a keyboard or second joystick) is needed to control walking. Alternatively, additional DoF may be controlled autonomously by the computer, for example through animations. Interestingly, computer-controlled animations do not necessarily reduce the perception of agency: for example, when playing Super Mario Bros,[5] the user only controls the overall movement and head direction of the virtual character, but the arm and leg movements are computer animated.

Another important point to consider is that the raw signals, measured by the sensors on input devices, need to be processed in order to enable good control. Firstly, input data is almost always filtered, for example with a low-pass or Kalman filter, in order to reduce noise. Further, if we consider the case of controlling the position of a User Representation with an input device, the most common operation is scaling the input values, such as in a control-to-display (C:D) gain. For example, moving the mouse by one centimeter might move the mouse cursor by many centimeters. Another important decision is whether input values are integrated over time, when controlling the state of the User Representation. With no integration, the position of the input device would directly control the position of the User Representation, such as when the position of a pen on a drawing tablet controls the position of the cursor. Joystick angles are usually integrated once, in order to enable rate (or velocity) control of a cursor or avatar. More rarely, the input signal is integrated twice in order to yield acceleration control (or "ice physics"), for example in the game Asteroids.[6]

More complex mapping operations can be used, such as including a "dead zone" around the center position of a joystick, where the signal is set to zero. Another common example entails nonlinear transfer functions, such as Pointer Acceleration functions, which make the C:D gain depend on the speed with which the mouse is moved (i.e., the cursor will move further if the mouse is moved faster). Such dynamical manipulations of the C:D gain may also be leveraged to induce pseudo-haptic illusions during interaction with virtual content. Users can thereby be tricked into perceiving haptic properties (i.e., friction, mass, stiffness, etc.) simply by modifying the visual feedback provided by the User Representation (see the review by Lécuyer, 2009). For instance, people perceive higher friction of a virtual space if the C:D gain is reduced while moving the User Representation across it. In mixed reality, the manipulation of visual movement of a handheld

virtual object modulates the user's perception of its weight and can reduce fatigue (Samad, Gatti, Hermes, Benko, & Parise, 2019; Taima, Ban, Narumi, Tanikawa, & Hirose, 2014). These weight illusions can be further strengthened through modification of the User Representation's visual appearance and based on the inclusion of biofeedback. In a system from Hirao and Kawai (2018), users are provided with visual feedback of their heart rate, and the virtual hands are shown trembling and turning red with the aim of reducing fatigue during the lifting of objects.

Similarly, dynamic manipulations of the C:D gain have been exploited for introducing flexible haptic feedback in immersive VR applications. In haptic retargeting, a physical object placed at a fixed location in the user's reaching space is used to provide haptic feedback from virtual objects placed in a range of different locations (Azmandian, Hancock, Benko, Ofek, & Wilson, 2016). This is done by introducing a dynamic spatial decoupling between the input device (e.g., a tracked hand) and the User Representation (e.g., a virtual hand) during the reaching action. Thus, in order to reach the virtual object with the virtual hand, the user has to redirect the reaching action so that the real hand arrives at the location of the physical proxy. Interestingly, in immersive VR applications, the redirection is unnoticeable for a wide range of spatial offsets (Cheng, Ofek, Holz, Benko, & Wilson, 2017). More subtle visuo-haptic illusions could be achieved via ad-hoc visual deformation of the User Representation, as when modulating the distance between the fingers of a virtual hand to an induce illusory perception of the size of objects that are held (Ban, Narumi, Tanikawa, & Hirose, 2013).

Dynamic changes of the C:D ratio can also be leveraged to support flexible interaction techniques, allowing users to manipulate virtual objects that are out of reach, as in the Go-Go interaction technique (Poupyrev, Billinghurst, Weghorst, & Ichikawa, 1996). With Go-Go, the user can naturally reach for close objects, with the virtual hand following their own hand movement in a 1:1 mapping (linear function) while close to the body. However, when extending the hand beyond a defined distance threshold to reach for far-away objects, the position of the virtual hand is computed as a non-linear function based on the distance of the user's physical hand to their body. This projects the virtual hand successively further away than the user's real hand, effectively extending the user's reach to select distant virtual objects. This technique was later adapted to enable control of distant physical elements in an actuated environment through an AR application (Feuchtner & Müller, 2017). In this work, the User Representation was a virtual arm that could stretch to twice the user's natural arm length, for example to manipulate a table's height or open a curtain. A further variant of Go-Go was implemented in combination with a wearable elastic armature which provides passive haptic feedback, depending on the degree of arm stretch (Achibet, Girard, Talvas, Marchal, & Lecuyer, 2015). By dynamically adapting the C:D gain to achieve an intended tension of the elastic band, the authors could simulate touching surfaces of varying stiffness, or varying the level of effort needed to move virtual objects.

The technical details of the sensorimotor mapping from user movements to the User Representation are critical to the user's experience of control during interaction. This sense of control, or agency, can vary according to the latency and sensorimotor pattern involved in the interaction (see *Sense of agency* Section 3.3), as well as the specific type of input device used (e.g., mouse, keyboard, or body tracking sensors) (Limerick et al., 2014) and the mapping functions that are applied. While spatial C:D manipulations (static or dynamic) with a coherent temporal control preserve the user's sense of agency and thus identification with the User Representation, temporal incongruences are tolerated to a lesser extent (Asai, 2015). In particular, the perception of control of a User Representation is critically influenced by how well users can predict the movement resulting from their input. Overall, input methods which lead to stronger feelings of agency over User Representations, may result in an enhanced user experience and more effective control.

3.5. *Spatial perception and tool embodiment*

Physical and virtual tools enable users to carry out actions in distant or virtual spaces that would otherwise be unreachable. A person can use a physical wooden stick to reach an object that is far

away, or a virtual cursor to select icons during indirect interaction. Remarkably, during such tool-use we are able to transfer the motor control of a specific body part to a mechanical or virtual end-effector, without significantly affecting our performance during the task at hand. End-effector, in this context, refers to the part of the tool that comes into contact with the object we are manipulating (Arbib, Bonaiuto, Jacobs, & Frey, 2009). A large amount of evidence has shown that this mastery in tool-use may be explained by the incorporation of tools into the mental representation of our body and the space immediately surrounding it, also known as peripersonal space (Berti & Frassinetti, 2000; Maravita & Iriki, 2004; Martel, Cardinali, Roy, & Farnè, 2016).

This peripersonal space is often conceptualized as the space that is within a reachable distance and that we can act upon. On a neural level, it can be defined as the space surrounding the physical body that is represented through multimodal populations of neurons which simultaneously respond to and integrate somatosensory, visual, and/or auditory information (Di Pellegrino & Làdavas, 2015). Importantly, this multimodal representation is coded in body-part-centric coordinates. For example, a visual stimulus that is encoded by neurons with tactile receptive fields on the person's right hand, will respond to visual stimuli close to the same hand, irrespectively of where in space the hand is located (Graziano, Yap, & Gross, 1994; Meredith & Stein, 1986). The space beyond our reach is called the extrapersonal space and is characterized by a different neural coding.

Iriki, Tanaka, and Iwamura (1996) provided the first evidence of active tool-use modifying the neural representation of the peripersonal space. In their study they showed that after using a tool to grab pieces of food, monkeys' visual receptive fields of bimodal neurons, which coded for touch and vision on the hand, were expanded to also include the tool (Figure 5A). Since then, several studies have shown that active tool-use can result in a remapping of the peripersonal space, in order to encompass the new space that is accessible and may be controlled with the tool (Martel et al., 2016).

Studies have further shown that tool use can affect distance perception and motor planning. Holding an elongated tool to reach a target, reduces the perceived distance between our bodies to this target. No such perceptual changes arise when only passively holding the tool, without an intention to act (Witt, Proffitt, & Epstein, 2005). Cardinali et al. (2009) found that the kinematics (i.e., movement patterns) of reach-to-grasp actions can change after using a tool to execute a similar action. The manner in which tools are mapped into the brain's spatial representations is also strongly influenced by the method used to control the tool. After tool-use, peripersonal space is reshaped to only include those regions of the tool that are relevant for manipulation. For example, when only the tip of the tool is used to execute an action, the remapping then occurs only around the tool's tip, and not along its handle. By contrast, if carrying out the task required manipulating an object with the middle, or the shaft of the tool (e.g., handle), then the remapping would happen along these areas of the tool. Thus, the functional locus of the tool is crucial for determining the changes in spatial configurations that will arise as a result of tool-use (Holmes, Calvert, & Spence, 2004).
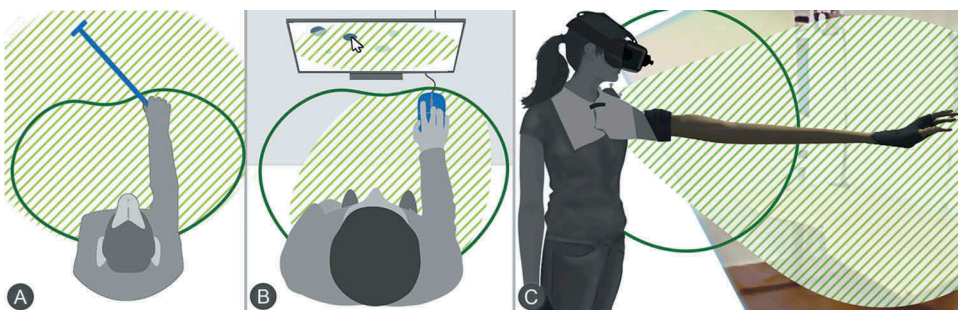

Figure 5. Peripersonal space is remapped, as a consequence of (virtual) tool use. In these examples, the dark green outline shows the peripersonal space before tool use, while the light green shadow illustrates its remapping as a result of: (A) a monkey using a long tool to reach food, (B) a user controlling a cursor through a computer mouse, and (C) experiencing the illusion of owning a long virtual arm in augmented reality.

Similar to physical tool use, changes in the peripersonal space can also arise as a result of "virtual" or "tele-operated" tool manipulations (Rallis, Fercho, Bosch, & Baugh, 2018). When playing a video game, users may perceive that the avatar acts as an extension of themselves, allowing them to perform actions in a virtual world. Such perceptions have a neural basis and seem to be modulated by the brain's spatial representations. For example, the visual receptive fields of bimodal neurons that code for a monkey's hand, are modified after the monkey learns to recognize his own actions on a video monitor (Iriki, Tanaka, Obayashi, & Iwamura, 2001). When the monkey views itself on a screen, the visual receptive fields of bimodal neurons are not only mapped to the monkey's hand, but also to the video screen where the hand is being represented. This provides evidence that remapping can take place, even when the space between the representations (e.g., screen) and where the actions happen (e.g., a person's or monkey's hand) are not physically connected, and may even exist in different "realities" (physical vs. virtual).

Reconfigurations of the peripersonal space also occur for indirect interactions with computers, where the input and output space are always separated, as in the case of executing actions on a screen with a computer mouse (Figure 5B). For example, it has been found that, after active mouse manipulations, the auditory peripersonal space is expanded to also include the region around the computer screen, although it is far from and disconnected from the body (Bassolino, Serino, Ubaldi, & Làdavas, 2010). This is typically characterized by an enhancement of the integration of auditory and tactile stimuli presented near to the body, compared to when it is presented far from the body (Farnè & Làdavas, 2002; Meredith & Stein, 1986). Such results are also found when participants passively hold the mouse with the hand they customarily use to control such a device. This reveals that the duration of the remapping can be long-lasting, which is probably due to training-induced brain plasticity (Green & Bavelier, 2008). Accordingly, it seems probable that interactions in immersive VR also lead to adaptations of the peripersonal space to the virtual tools that the user manipulates in the virtual environment. This may involve handling distant objects with a long virtual tool, or having an elongated arm that extends to allow access to objects beyond reach (Figure 5C) (Feuchtner & Müller, 2017; Kilteni, Normand, Sanchez-Vives, & Slater, 2012).

In HCI, Bergstrom-Lehtovirta et al. (2019) have recently adapted a visuo-tactile interference paradigm to measure tool extension when using different User Representations in VR. They found an enhancement in the integration of visual and tactile information when the user was represented in VR with realistic avatar hands, compared to a pointer. Along the same lines, Alzayat et al. (2019) have also lately proposed a method based on attentional processing to assess tool embodiment (i.e. feel that the tool becomes an extension of one's own body) for different input methods. They carried out a VR experiment, where users handle a virtual tool either by using its physical counterpart, by direct hand movements, or through controllers. Their results indicate that users pay more attention to the virtual task than to the virtual tool, when the tool is working correctly as opposed to being defective. They also found that users tend to pay more attention to the task when using direct hand inputs, compared to the controllers. Based on the results obtained, the authors conclude that the Locus-of-Attention Index presented in the paper, seems to be a promising measure for tool embodiment when testing different devices and input methods.

Finally, another interesting aspect is that when threatening stimuli or warning signals are presented in the peripersonal space, they are treated by the brain as more behaviorally relevant, and thus demand more attention, compared to when presented in the extrapersonal space (Ho & Spence, 2009; Previc, 2000). Based on this knowledge, we could also speculate that the space where notifications, alarms, or feedback is presented with respect to the position of the User Representation, can impact differently on user experiences and capture their attention.

3.6. *Visual perspective and sense of self-location*

In two-dimensional (2D) space (e.g., side-scroller games), users perceive their avatar from 3PP. However, 3D environments allow users to experience their User Representation from both 3PP or
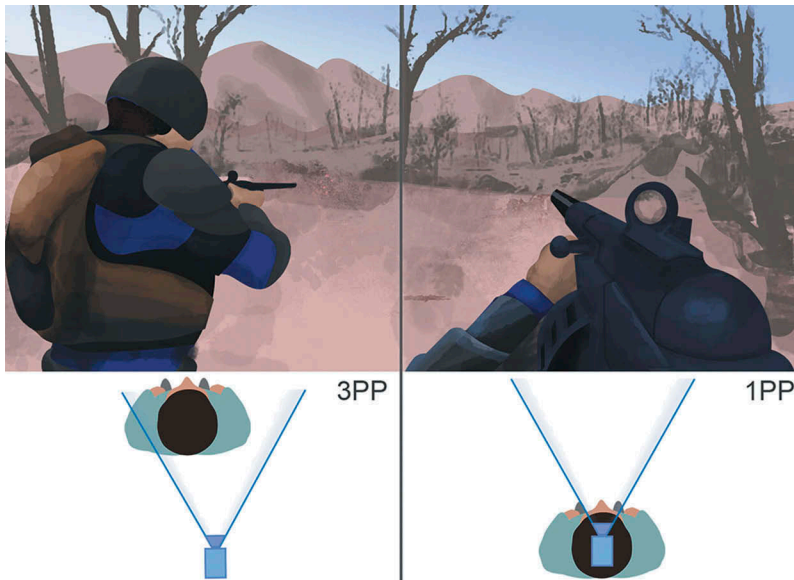
Figure 6. When interacting in virtual 3D space, the User Representation is either perceived from First-Person Perspective (1PP) or Third-Person Perspective (3PP). This is particularly the case in computer games, where the user may follow an avatar from behind, resulting in 3PP (left), or may experience the environment through the eyes of the avatar (right), thus resulting in 1PP.

1PP (Figure 6). In 1PP, the camera is placed in the same location as the User Representation, so the player has the sense of exploring the digital environment directly from the representation's visual perspective. In those cases, users can literally experience the environment from the "eyes" of the virtual avatar, which may lead to strong feelings of immersion and presence (Denisova & Cairns, 2015). However, contradictory evidence has been found in this regard. For example, when experiencing a role-playing game on a TV screen, it has been reported that 1PP leads to a higher level of immersion, compared to 3PP (Denisova & Cairns, 2015). However, in immersive VR, there is no evidence of differences in the feelings of presence between 1PP and 3PP (Gorisse, Christmann, Amato, & Richir, 2017). It may be possible that the interaction mode plays a role in the experienced level of presence, irrespective of the visual perspective of a User Representation. The sense of "presence", also known as place-illusion or sense of being inside the place depicted in VR, seems to arise as a result of the interface supporting sensory and motor behaviors that parallel those of physical reality (Slater, 2009). Based on this notion, immersive VR provides interactions which more closely resemble the ways humans perceive the real world, when compared to regular monitors or TV screens. For example, HMDs allow users to explore a virtual environment through direct head movements. These sensorimotor contingencies can result in high levels of presence when using a VR system, independently of having a 1PP or 3PP. However, when interacting through a controller and a observing the User Representation on a TV screen, the role of 1PP in accomplishing a high degree of presence or immersion might be more prominent.

3PP allows users to view their representations from a distance, often from behind, from the side, or from above. In these cases, users can access informative cues about the spatial configuration and events occurring in the virtual environment. For this reason, applications with the main goal of navigation and wayfinding typically benefit from giving users a 3PP. In fact, it has been shown that 3D map navigation from 3PP results in shorter orientation times, better clarity ratings, lower workload, and higher preference scores, compared to 1PP (Burigat, Chittaro, & Sioni, 2017). A VR study supports this view further, since it was found that 3PP improves the users' spatial awareness and perception of the virtual environment surrounding the avatar, compared to 1PP (Gorisse et al., 2017).

The use of HMDs to give users a 1PP view of a life-size virtual avatar can even result in body ownership illusions, i.e. perceiving a virtual body as being one's own real body (for details, see Section 3.7 *Feeling of body ownership*). Substantial amounts of research have demonstrated that 1PP is a crucial aspect of inducing these types of illusions regarding a virtual artificial body seen in VR (Maselli & Slater, 2014). A further important concept related to visual perspective and camera position is that of self-location. This term refers to the perceptual experience of occupying a given location in the environment (Blanke & Metzinger, 2009; Serino et al., 2013). While during body ownership illusions self-location corresponds to the location of the real body (occluded by the immersive device), alterations to the sense of self-location can also be induced when providing a 3PP over a virtual body, as during out-of-body illusions (Bourdin, Barberia, Oliva, Slater, & Passingham, 2017). Depending on the modality in which visuo-tactile or other multimodal stimulations are delivered, the illusory changes in the perceived self-location may vary, resulting in either egocentric (i.e., sense of location in reference to one's own physical body or "self") (Ionta et al., 2011; Lenggenhager, Tadi, Metzinger, & Blanke, 2007) or allocentric (i.e., sense of location relative to the world or the environment) perceived drifts in self-location (Maselli, 2015).

In egocentric drifts toward an artificial body, participants report the sensation that a virtual body that is located in front of them is their own body (Ionta et al., 2011; Lenggenhager et al., 2007). Moreover, they also mislocalized themselves toward the virtual body, in a position that is outside the location of their own physical body. This has been reported when participants see a life-size avatar from a 3PP through an HMD, and an experimenter strokes the participants' back while they see the avatar's back being touched with a virtual stick. However, the participants do not see a visual counterpart of the felt touch on their own back (i.e., a stick touching their back) and there is no perfect spatial match between the felt and seen touch on the avatar's back. The temporal synchronicity of the touch, and the fact that the feedback is seen in the same body part as the felt one, is sufficient to induce perceived changes in self-location. A recent study has also shown that it is possible to experience a feeling of body ownership of anavatar seen from 3PP in VR, but only if users are able to control the avatar's motion through their own body movements, or when congruent tactile stimulation is provided (Galvan Debarba et al., 2017). However, body ownership was reported to be stronger in 1PP, and users preferred to have the possibility of switching visual perspectives.

The allocentric drift differs from the egocentric based mainly on the presence of specific environmental landmarks. Here, the person experiences what is known as an out-of-body illusion, described as the feeling of the "self" or "center of awareness" being located outside one's own physical body, and having the sense that one is seeing his/her own body from another person's perspective (Guterstam, Björnsdotter, Gentile, & Ehrsson, 2015). This illusion has been induced by filming a real body from a different position within the experimentation room, so that the participants see their own back through an HMD, as if they were sitting behind themselves (Ehrsson, 2007). Using a rod, the experimenter provides tactile feedback on the participant's chest and on the camera, so the participants feel as if they were being touched in the same location. These types of manipulation trigger an allocentric remapping of self-location, where the person feels that he is located outside of his/her own body in a given environment/room. Similar out-of-body experiences have also been simulated in VR (Bourdin et al., 2017).

3.7. *Feeling of body ownership*

Body ownership refers to the feeling that we have and own a corporeal manifestation in the physical world through which we can experience and manipulate our environment. The conviction that our hands are in fact our own and part of our body, is an example of body ownership. Research in cognitive neuroscience has demonstrated that we can also experience this feeling toward an artificial body, which is then known as a body ownership illusion (Kilteni, Maselli, Kording, & Slater, 2015). This type of illusion can be induced over a rubber hand (Botvinick & Cohen, 1998), a mannequin body (Petkova & Ehrsson, 2008), a virtual hand (Slater, Pérez Marcos, Ehrsson, & Sanchez-Vives,

2009), an entire virtual body (Maselli & Slater, 2013), a virtual mirror reflection (Gonzalez-Franco et al., 2010), and even a virtual body that differs radically from the participant's actual physical appearance (Banakou, Groten, & Slater, 2013; Seinfeld et al., 2018). This malleability of our mental body representation may enable us to create radically new types of User Representations and related interaction techniques. Instead of being represented by abstract virtual objects experienced from 3PP, users can see virtual bodies from a 1PP, and perceive them as their "own".

Body ownership illusions are effectively induced through congruent multisensory stimulation. In the rubber hand illusion, a participant's real hand is hidden from view and replaced with a hand-like object (Figure 7A). When synchronous tactile stimulation is provided to both the real and fake hands, so that participants see the fake hand being stroked while feeling congruent touch on their real hand, they typically experience the illusion that the rubber hand is part of their own body (Figure 7B). At this point, if somebody attempts to harm the rubber hand, participants will react as if the threat was addressed to their own real body (Ehrsson, Spence, & Passingham, 2004; González-Franco, Peck, Rodríguez-Fornells, & Slater, 2014). Apart from integration of visual and tactile stimuli, body ownership can also be elicited through visuo-proprioceptive (Maselli & Slater, 2013) or visuo-motor stimulation (Sanchez-Vives, Spanlang, Frisoli, Bergamasco, & Slater, 2010; Tsakiris, Prabhu, & Haggard, 2006).

VR is a powerful tool for inducing body ownership illusions. With this technology, 1PP of a collocated (immobile) virtual body is sufficient to evoke body ownership, due to the important role played by visual and proprioceptive information in such illusions (Ferri, Chiarelli, Merla, Gallese, & Costantini, 2013; Maselli & Slater, 2013). The inclusion of additional congruent visuo-tactile and/or visuo-motor stimulation can further strengthen or weaken the illusion, depending on the associated spatio-temporal congruency of the stimulation (Kilteni et al., 2015; Kokkinara & Slater, 2014; Maselli & Slater, 2013). When participants see a virtual sphere touching their virtual hand and simultaneously perceive touch at the same location on their own real hand (e.g., someone strokes the person's hand with a real brush or vibration of a motor; Figure 8A), they have the illusion of being touched by a virtual sphere (Slater et al., 2009) (Figure 8A). Moreover, exposure to visuo-motor contingencies through real-time body tracking typically leads to an even stronger feeling of ownership of the artificial body (Kokkinara & Slater, 2014; Tsakiris et al., 2006) (Figure 8B).
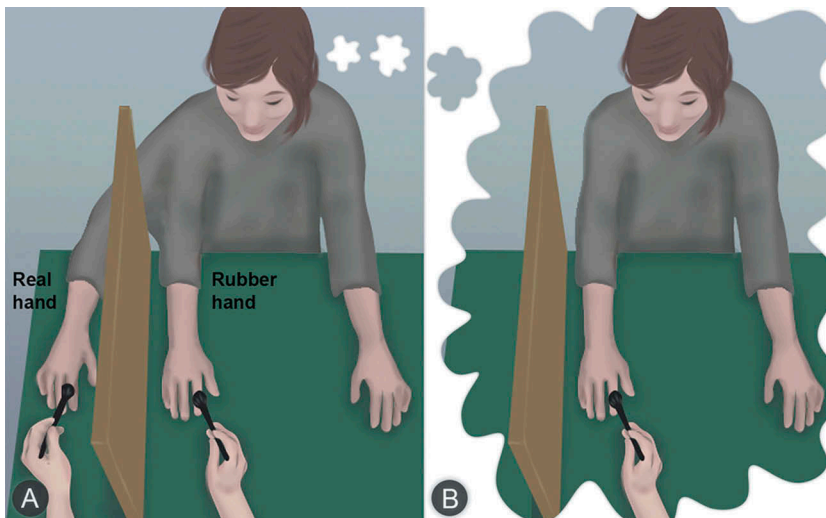

Figure 7. In the rubber hand illusion, synchronous tactile stimulation is applied simultaneously to the rubber hand and the participant's hidden real hand (A). Seeing the rubber hand being stroked while feeling congruent strokes on the own hand gives the participant the illusion that she can 'feel' with the rubber hand. The rubber hand is thereby attributed to the own body, replacing the real hand (B).
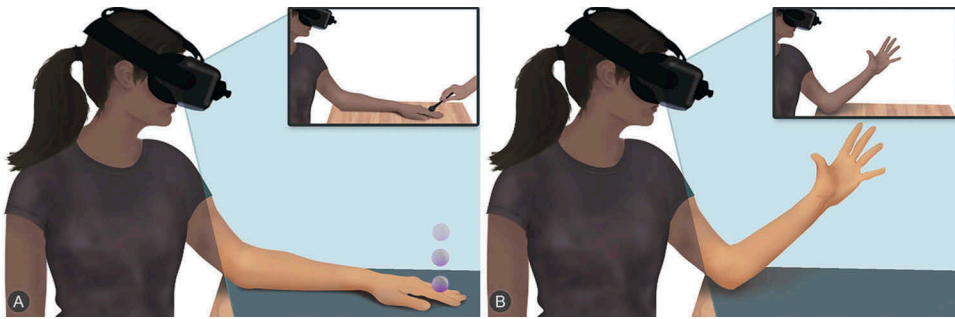
Figure 8. Seeing a virtual hand in place of our own can lead to illusory body ownership of the virtual hand (aka. virtual hand illusion). As in the RHI, this illusion can be elicited through synchronous visuo-tactile stimulation, e.g., (A) the participant sees a ball bouncing on the virtual hand, while in reality the experimenter supplies synchronous tactile stimuli by tapping the hand with a wand (inset image shows real world). Motion tracking technology further allows visuo-motor synchrony: (B) observing the virtual hand mimic the motions of their own hand (inset image shows real world).

In addition, when users see a virtual body moving in accordance with their own real movement, a strong sense of agency over the virtual body is evoked. The relation between body ownership and agency has been investigated in many studies. Although evidence suggests that both phenomena are fully dissociable (Kalckert & Ehrsson, 2014), there seem to be reciprocal effects between both sensations (Braun et al., 2018; Tsakiris, Schütz-Bosbach, & Gallagher, 2007). For example, experiencing agency strongly enhances feelings of body ownership. Surprisingly, the perception of body ownership can also make participants misattribute the authorship of an action to themselves, which is strongly related to vicarious agency, referring to feelings of authorship for the actions of others (Wegner et al., 2004). This can occur when users see an artificially triggered action (e.g., opening or closing of a hand) performed by the virtual body they are embodying (Banakou & Slater, 2014; Kokkinara et al., 2016). The section *Sense of agency* provides detailed descriptions of studies, which found that participants can misattribute to themselves the action of speaking and walking, when they experience body ownership of thevirtual body carrying out these actions.

In addition to bottom-up multisensory and sensorimotor contingencies, body ownership illusions are also strongly modulated by top-down factors, such as the visual appearance of the artificial body. When the shape of an artificial object does not satisfy anthropomorphic constraints, the illusion is completely inhibited, such as when regarding wooden blocks, abstract virtual cursors, or sticks (Kilteni et al., 2015; Tsakiris & Haggard, 2005; Yuan & Steed, 2010). Neither does the illusion occur when the artificial body is placed in an anatomically impossible configuration, such as when an artificial hand is seen rotated by 90º or 180º degrees (Ehrsson et al., 2004; Maselli & Slater, 2013). Unrealistic visual appearance, such as the visual discontinuity of the artificial body, also reduces the feeling of body ownership. It has been found that observation of a static virtual hand that is visibly disconnected from the rest of the virtual body or arm, results in weaker feelings of body ownership and decreased physiological reactions to a virtual threat, compared to fully connected virtual hands (Tieri, Tidoni, Pavone, & Aglioti, 2015).

Evidence suggests that for unrealistic objects, which to some degree retain anthropomorphic anatomical features, it is still possible to induce body ownership illusions by delivering congruent dynamic multimodal stimulation to the users. This is the case for mannequin bodies (Petkova & Ehrsson, 2008), elongated arms (Feuchtner & Müller, 2017; Kilteni et al., 2012), non-collocated limbs (Feuchtner & Müller, 2018), supernumerary limbs (Bashford & Mehring, 2016; Guterstam, Petkova, & Ehrsson, 2011), and even for bodies that are mostly invisible of which only the extremities (hands and feet) are seen moving in synchrony with the user (Kondo et al., 2018). In summary, converging evidence from over twenty years of experimental research suggests that it is possible to induce the illusion of body ownership of a bodily-shaped external object – and therefore also of a virtual User

Representation – by providing subjects with a 1PP of the fake body and exposing them to different combinations of congruent multimodal contingencies. These can range from static visuo-proprioceptive stimuli to more complex visuo-motor or visuo-tactile correlations, whereby modulations are associated with the visual appearance of the object (Kilteni et al., 2015; Maselli & Slater, 2013).

Visual fidelity of the artificial body was also shown to significantly modulate the processing of sensory inputs (i.e., visuo-haptic or visuo-auditory), the sense of presence, and the perception of objects within the virtual environment. Ogawa et al. (2019) found that the size of the virtual hands only impacted on the estimated size of virtual objects when highly realistic virtual hands were used for interaction, but not when controlling iconic-abstract hands. Similar results were found by Jung, Bruder, Wisniewski, Sandor, and Hughes (2018), who reported stronger body ownership and spatial presence, and more accurate size estimates of virtual objects, for participants who were given high-fidelity personalized virtual hands. Also comparing hands of different sizes, Lin et al. (2019) did not observe clear differences in body ownership, but found that the input modality played an important role. Users preferred systems that directly enabled them to control the virtual hand with their own hand motions, instead of hand-held devices. With regard to the perception of multiple sensory inputs, Schwind, Lin, Di Luca, Jörg, and Hillis (2018) found that the appearance of virtual hands modulates visuo-haptic integration when perceiving surface irregularities. Further, Tajadura-Jiménez, Banakou, Bianchi-Berthouze, and Slater (2017) found that congruency between a virtual body (i.e., embodiment in a child avatar) and auditory cues (i.e., distorting the user's voice to match a child's voice) related to that body, leads to stronger body ownership.

Overall, most evidence has shown that it is not possible to experience body ownership over external non-antropomorphic objects. However, some authors claim that is possible to feel body ownership for non-corporeal objects (e.g., abstract virtual shapes, mobile phones, physical tools), at least to some extent (Liepelt, Dolk, & Hommel, 2017; Ma & Hommel, 2015). Yet, body ownership scores in these cases are always lower for non-corporeal objects than those reported for bodily-shaped objects. It is possible that such results can be explained by other factors not linked to body ownership, such as agency or tool embodiment. However, these explanations remain to be verified by future studies. Interestingly, some research suggests the possibility of experiencing body ownership for non-humanoid avatars, such as animal avatars. Oyanagi and Ohmura (2017) observed that correspondent visuo-motor mapping between a life-size bird avatar and a person's real body, led to a strong illusion of body ownership. In an earlier study, Ahn et al. (2016) showed that these body transfer illusions for animalistic avatars could be leveraged to increase involvement with nature and an awareness of environmental risk. Based on the aforementioned studies, it is possible that users are more prone to experiencing body ownership of living organisms, compared to inanimate objects. This topic should be further researched.

Interestingly, the experience of body ownership has been found to affect perceptual processing. For example, participants' sensitivity in detecting tactile stimuli was shown to decrease during body ownership illusions (Folegatti, de Vignemont, Pavani, Rossetti, & Farnè, 2009; Zeller, Friston, & Classen, 2016; Zopf, Harris, & Williams, 2011). Moreover, illusory body ownership can relax temporal constraints for multisensory integration (Maselli, Kilteni, López-Moliner, & Slater, 2016) When visual and tactile stimuli are presented on the participant's real hand and on a collocated virtual hand during a body ownership illusion, participants are unable to detect temporal delays between these two types of stimulation that they would otherwise be able to notice. This is interpreted as a consequence of tactile and visual stimuli being attributed to the same origin, i.e. one's own body. A further study found that users may adapt their movements to avoid harming their virtual hand, if this hand representation looks realistic, which in turn elicits body ownership thereof (Argelaguet, Hoyet, Trico, & Lecuyer, 2016). These results have important implications for HCI, as specific applications might benefit from inducing body ownership over a User Representation. For example, to mitigate the effects of temporal delays or latencies of devices that provide visuo-tactile or visuo-motor feedback.

In regards to HCI, one of the important aspects of the body ownership illusion lies in its potential to not only modify the users' experience, but also their perceptions, attitudes, and behaviors, which can be modulated according to the visual features and semantic properties of the owned virtual body. Yee and Bailenson (2007) first discovered that participant behavior and attitudes in VR could be differently modulated, depending on the appearance of their digital self-representation. For instance, they found that participants who were assigned more attractive and taller avatars were more confident and displayed higher intimacy, compared to those embodied in less attractive and shorter avatars. The authors refer to this phenomenon as the Proteus Effect. Similar effects were shown by Banakou et al. (2013) where experiencing body ownership toward a virtual child body subsequently led participants to overestimate the size of objects and associate more easily with infancy-related concepts, compared to having the body of an adult avatar. The appearance of the owned body can further modulate pain perception (Matamala-Gomez, Diaz Gonzalez, Slater, & Sanchez-Vives, 2018; Nierula, Martini, Matamala-Gomez, Slater, & Sanchez-Vives, 2017). Kilteni, Bergstrom, and Slater (2013) observed that participants who were asked to play drums in VR exhibited significantly different movement patterns when their User Representation was dark-skinned and casually dressed, compared to light-skinned and formally dressed. Furthermore, it has been shown that the embodiment of a dark-skinned body can lead to a reduction in implicit racial bias (Peck, Seinfeld, Aglioti, & Slater, 2013). Similarly, another study found that domestic violence offenders improve their emotional recognition skills after experiencing the illusion of being in the body of a female victim (Seinfeld et al., 2018). Together, these results indicate that when a user experiences body ownership of a virtual User Representation, the semantic properties of this representation can play an important role in shaping how the user will behave and perceive (Maister, Slater, Sanchez-Vives, & Tsakiris, 2014).

4. Examples

In the following section we will discuss the concepts presented in this work within the context of two particular computer-supported activities: drawing (example A) and playing a game of tennis (example B). Further, Figure 9 provides an illustration of these concepts, involving different types of interactive systems and various User Representations.

Example A: Drawing

Computer-supported drawing is a common activity, both in a professional capacity (e.g., sketching new ideas or plotting graphs by hand), as well as for enjoyment. Desktop applications such as Microsoft Paint[7] or Adobe Photoshop[8] enable a user to draw on a virtual 2D canvas depicted on-screen with their User Representation, in this case a mouse cursor.

In applications such as Google Tiltbrush,[9] Gravity Sketch,[10] and Leap Motion Paint,[11] users can draw shapes in mid-air with virtual representations of hand-held controllers or iconic virtual hand representations. In the case of virtual hands, motion tracking technology enables faithfully recreating movements of the user's real hands.

A common aspect of these applications is the inclusion of User Representations that dynamically change their **visual appearance** to signify the use of different tools (see Section 3.1. *Visual appearance and affordances* for a detailed literature review). By relying on users' previous knowledge of physical tool-use, the User Representation leverages visual cues, semiotics, and functional roles attributed to these tools for direct manipulation (Beaudouin-Lafon, 2000; De

---

[7]https://en.wikipedia.org/wiki/Microsoft_Paint.
[8]https://www.adobe.com/at/products/photoshop.html.
[9]https://www.tiltbrush.com/.
[10]https://www.gravitysketch.com/.
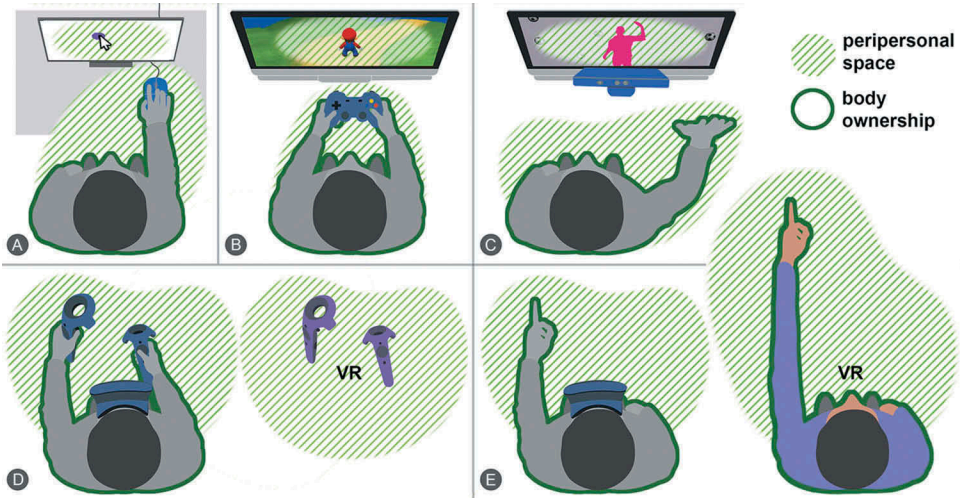[11]https://gallery.leapmotion.com/paint/.

Figure 9. We illustrate the application of our framework with 5 examples, showing how peripersonal space, body ownership, and visual perspective can change, depending on the interface and input devices used. (A) Mouse and pointer, (B) Gamepad and 3D avatar (3PP), (C) Kinect and mirror image, (D) VR HMD with motion tracking and humanoid avatar with elongated arm (1PP).

Souza, 2005; Kato et al., 2000; Norman, 2008). In 2D applications, the cursor frequently takes the shape of a physical tool used in the real world to perform an analogous task, such as a virtual pencil, brush, or eraser tool tip in Photoshop or Paint. In the case of TiltBrush, Gravity Sketch, and Leap Motion Paint, the use of visual cues to represent tools is more restricted and abstract. Despite the tip of the virtual controllers in TiltBrush slightly changing its visual appearance to signify the action of drawing, the representation does not radically change its visual properties when selecting a spray can or paintbrush, etc. Thus, a user may not be aware of what tool is currently selected, until they begin to draw. In Leap Motion Paint, the tip of the virtual index finger reflects which color is selected from a given color palette, and otherwise offers a more restricted selection of virtual tools limited to painting and erasing. Gravity Sketch seems to overcome some of these shortcomings by enabling the User Representation to change from a controller to a specific tool. However, to the best of our knowledge, there is no 3D drawing application showing virtual hands directly holding a virtual tool for drawing, or where the representation changes in real-time to realistically visualize the specific selected painting tool. In the future, it would be useful to address how the level of abstraction of User Representations impacts learning and task-performance in mixed reality systems.

Between the examples of drawing applications presented here, it is likely that **spatial perception** and **tool embodiment** are also differently impacted (see Section 3.5. *Spatial perception and tool embodiment* for a review of the literature). Research suggests that 2D desktop applications such as Photoshop and Paint, where users are represented through a mouse cursor, lead to a remapping of peripersonal space toward the computer screen where the mouse cursor is presented (Bassolino et al., 2010; Gozli & Brown, 2011). Due to the spatial collocation with the real user's body in the case of virtual controllers or hands in VR, the remapping of peripersonal space will most likely occur along the tip of these virtual tools (Holmes et al., 2004; Rallis et al., 2018). However, studies in VR suggest that tool embodiment will be stronger for a virtual hand compared to the controller representations (Bergstrom-Lehtovirta et al., 2019). Whether such phenomena might impact user experience and performance when drawing in either Google Tilt Brush or Leap Motion, is a topic that also requires future research.

With regard to **visual perspective**, desktop applications (i.e., Photoshop and Paint) provide a 3PP of the cursor, while immersive VR painting applications present the controller (i.e., TiltBrush)

and virtual hands (i.e., Leap Motion Paint) from 1PP. Research indicates that users may have an increased spatial awareness when painting with a User Representation seen from 3PP (Burigat et al., 2017; Gorisse et al., 2017). On the other hand, 1PP has been shown to lead to higher levels of immersion, which may also be the case when painting in VR, compared to painting on a traditional computer screen (Denisova & Cairns, 2015). In any case, evidence suggests that in an immersive application it is advisable to allow the user to switch visual perspective as desired, since at different stages of the task they might desire a broader view of the painting (i.e., 3PP) or wish to feel more immersed (i.e., 1PP) (Debarba et al., 2017; see Section 3.6. *Visual perspective and sense of self-location*).

Furthermore, based on the reviewed literature, it is highly likely that users will experience **body ownership** of the virtual hands in the Leap Motion paint application (see detailed explanations in Section 3.7. *Feeling of body ownership*). This is supported by an ample amount of evidence, showing that body ownership illusions can be evoked even for unrealistic, iconic hands, simply due to their anthropomorphic characteristics and 1PP (Lin & Jörg, 2016; Ogawa et al., 2019). However, based on past studies carried out in passive visuo-proprioceptively congruent conditions, the fact that the virtual hands are disconnected instead of being part of a full virtual body might weaken the perceived body ownership (Tieri et al., 2015). The iconic nature of the virtual hands might also lead to a decrease in sense of presence (Jung et al., 2018), modulate the perception of textures (i.e., bumps and holes) based on pseudo-haptic illusions (Schwind et al., 2018), and lead to overestimation or underestimation of the size of virtual objects within the scene (Jung et al., 2018; Linkenauger et al., 2013; Ogawa et al., 2019). Further, studies on typing applications in VR have found that transparent virtual hands can better facilitate the visibility of keys while typing, compared to fully visible hands (Grubert et al., 2018). Such transparency might also potentially be exploited in painting applications, or at least be available as an option to the user.

In contrast to virtual hands, past research suggests that it is highly unlikely for users to experience body ownership of virtual controllers or a mouse cursor (Tsakiris, Carpenter, James, & Fotopoulou, 2010; Yuan & Steed, 2010). Nevertheless, the **sense of agency** is strong for all User Representations, independently of their visual appearance, since the user's movements are mapped into analogous movements of the User Representation in real time (Zopf et al., 2018). A summary of the basic principles underlying agency can be found in Section 3.3. *Sense of agency*. Desktop drawing applications traditionally support relative mouse input, whereby the 2D movement of the mouse on the horizontal desk is transformed into a 2D movement of the cursor on the vertical screen. Hereby the transfer function and C:D gain must be carefully considered. When using pointer-acceleration, the mouse gain increases with the movement speed, which might lead to a distorted drawing e.g., when tracing a circle with varying stroke speed. In addition, many such applications support direct input through interactive pen tablets (e.g., Wacom tablet[12]), which allow drawing on a touch surface with a finger or a stylus. While the **input-output mapping** with the mouse is relative, the drawing tablet supports absolute input, meaning that every position on the tablet is mapped to exactly one position on the canvas (1:1 mapping).

VR applications that allow drawing and painting in 3D space through controllers or hands, use position control with a 1:1 absolute mapping. This means that the User Representation is collocated with the user's physical hands or hand-held controllers. Moreover, the virtual user representation may closely reflect the users' hand movements in 3D space with 6 DoF. This is achieved by tracking the position and orientation of the controllers (e.g., HTC Vive controllers), or the position and joint angles of the user's hands (e.g., Leap Motion). Some benefits of virtual hands that reflect users' real hand motions are higher perceived realism, presence and body

---

[12]https://www.wacom.com/en-us/products/pen-tablets.

ownership, compared to control exerted through hand-held devices (Lin et al., 2019). A review of the literature on these topics can be found in Section 3.4. *Motor control, input methods, and mappings.*

In addition to continuous visuo-motor feedback through the User Representation and dynamic changes in its visual appearance, many graphics editors provide further **multimodal feedback**, e.g. haptic and acoustic. In immersive VR drawing applications, multimodal feedback is often more abundant, to support effective mid-air interaction (Cornelio-Martinez et al., 2017). For example, when performing a brush stroke with the virtual controller or hand, the visual feedback of the stroke appearing in mid-air is accompanied by spatial sound, potentially providing additional cues about the type of tool being used (e.g., sound of a stiff brush on canvas). In Leap Motion Paint, when dipping the fingers into color pots on the palm-menu, drip and splash sounds also tell the user that a new color has been selected. Controllers often provide haptic feedback, such as vibration, which increases the perceived realism during interaction with virtual menus and virtual tools. Importantly, such multimodal feedback is generally directly related to user-generated actions, potentially enhancing the sense of agency (Limerick et al., 2014) and task performance (Georgiou et al., 2018; Hwang, Son, & Kim, 2017), which is an aspect that future research should explore further. A detailed review of the impact of multimodal feedback with respect to User Representations, can be found in Section 3.2. *Multimodal feedback.*

Example B: game of tennis

Tennis has a very long tradition in HCI, with Tennis for Two[13] developed by the American physicist William Higinbotham in 1958, being one of the earliest video games. This was followed by Pong that portrays table tennis.[14] Console video games, such as Mario Tennis,[15] Topspin 4,[16] and Kinect Sports Tennis (Season 2),[17] also support playing tennis. More recently, tennis has also found its way into VR, as in First Person Tennis.[18] This is a fully immersive game, where users are provided with a 1PP experience of a tennis game simulation.

The **visual appearance** of User Representations included in tennis games has evolved from representing users with minimal virtual objects (i.e., lines as paddles) to more complex ones (i.e., avatars). From a semiotic or signifier standpoint (De Souza, 2005; Norman, 2008), being able to play tennis with a virtual humanoid representation would closely resemble how we play tennis in the real world (i.e., playing tennis with our own real body). However, avatars may feature additional signifiers that highlight their specific abilities, which in turn corresponds to how the avatar responds to the user inputs. For instance, in Mario Tennis small characters like Baby Mario or Yoshi are especially fast. Aiming for a more realistic simulation, TopSpin 4 allows the selection of avatars depicting professional tennis players (e.g., Nadal and Federer) and replicates their playing style. The fact that users play tennis with virtual representations of expert players, might actually result in improved performance. This hypothesis is supported by research on the Proteus Effect in VR, where users display more confidence when embodying taller avatars (Yee & Bailenson, 2007), or perform better in a cognitive task when embodied in an Einstein avatar (Banakou, Kishore, & Slater, 2018). However, such research is less conclusive when a user controls an avatar from 3PP, so that further research is needed (Ash, 2016).

With respect to **spatial perception** and **tool embodiment**, the effects are likely to be similar to what has already been explained in the drawing example. With respect to console-based games such as Pong, Mario Tennis, Kinect Tennis, and TopSpin 4, it is likely to observe a remapping of peripersonal space toward the screen where the avatars or virtual objects are

---

[13]https://www.bnl.gov/about/history/firstvideo.php.
[14]https://en.wikipedia.org/wiki/Pong.
[15]https://en.wikipedia.org/wiki/Mario_Tennis.
[16]https://en.wikipedia.org/wiki/Top_Spin_4.
[17]https://en.wikipedia.org/wiki/Kinect_Sports:_Season_Two.
[18]https://store.steampowered.com/app/454140/First_Person_Tennis__The_Real_Tennis_Simulator/.

controlled through an input device (Bassolino et al., 2010). By contrast, in VR games such as First Person Tennis, evidence supports the notion that the peripersonal space will be remapped to encompass the space immediately surrounding the virtual body (Noel et al., 2018; Serino et al., 2018). Since in First Person Tennis, the avatar is holding a virtual tennis racket, it is possible that peripersonal space is extended to include the virtual tool (i.e., tennis racket), analogous to how physical tools can be incorporated into the body schema (Berti & Frassinetti, 2000; Martel et al., 2016). This is also an interesting topic that requires further research.

In terms of **visual perspective**, the impact of playing tennis in 1PP compared to 3PP is again probably similar to the drawing example. Tennis video games on video consoles provide users with a 3PP (i.e., Pong, Mario Tennis, TopSpin 4 and Kinect Sports Tennis), where users have a broad view of the 3D tennis court and see their representation from behind or above. This wide view provided by 3PP can increase the awareness of spatial configurations and additional information of events happening within the scene, potentially improving playing strategies (Burigat et al., 2017; Gorisse et al., 2017). During gameplay with an HMD, as in First Person Tennis in VR, the user is endowed with a 1PP of a life-size avatar holding a racket. In this case, the sense of viewing the virtual world from the "eyes" of the avatar, may lead to strong feelings of immersion and presence (Denisova & Cairns, 2015).

Due to the fact that **body ownership** illusions can arise for virtual bodies with plausible morphology that are perceived from 1PP (Kilteni, Groten, & Slater, 2012), it may be possible to experience a body ownership illusion when playing First Person Tennis. On the other hand, research findings do not support body ownership arising in games like Mario Tennis, Kinect Sport Tennis, and Topspin 4, due to the 3PP and the lack of collocation of the user with the representation (Maselli & Slater, 2014), and furthermore, due to the mapping of controls .

Analogous to the drawing example, the **sense of agency** over the User Representation in these tennis games is accomplished through visuo-motor correlations (Zopf et al., 2018). When the user executes a motor action (i.e., movement of joystick, pressing a button, body motion inputs), there is an almost immediate change in the actions performed by the User Representation (i.e., walking, jumping, wiggling racket, etc.). However, it is interesting to note that in these games, several avatar actions are automatically animated. For instance, in Mario Tennis, the characters display automatic winning or losing animations. In Kinect Sports Tennis the user's real leg movements are not directly reflected by the avatar. Instead, the virtual character side-steps or walks forward and backwards in response to the player moving in the corresponding direction. Moreover, these games frequently include some methods that assist the user in reaching the target (i.e., a virtual tennis ball) and hitting it. Interestingly, here, the break of control does not seem to disrupt the user's perceived sense of agency over the avatars' actions while playing. This aspect may be related to expectations about the specific consequences of pressing a button (Haggard, 2017), but this requires further research.

As discussed, the **mapping of controls** can vary strongly depending on the **input method**. In Pong the mapping complexity is minimal: the user moves the paddle by turning a rotary knob. This rotation is mapped to a 1D movement of the paddle (up/down). Newer hand-held input devices typically include a multitude of buttons and joysticks and the mapping becomes increasingly more complicated, in order to support a wider range of actions. For instance, Mario Tennis allows moving the avatar by means of a rate-control joystick, and two main buttons trigger different shots. Topspin 4 is played with even more complex controllers that usually feature 2 joysticks, 10 or more buttons and 2 triggers. In Kinect Sports Tennis, the user's full body motion is tracked with a camera sensor and the avatar mimics the user's movements. To hit the ball with the racket, the user simply performs a swinging motion with his arm, which resembles the motions of an actual tennis game. Similarly, in First Person Tennis, the user's motions are tracked and the avatar mimics these movements. In such motion controlled games, the user may vary shots by holding the hand or controller at different angles, swinging at different speeds, or hitting the ball with different parts of the virtual tennis racket. Future studies should address in more

detail how perception and control are modified by the use of these different types of mappings and input systems. In relation **to multimodal feedback**, auditory and tactile cues contribute to creating a realistic environment and allow users to perceive that the interface is responsive to their actions (see Section 3.2. Multimodal feedback section; Obrist et al., 2016). For instance, users' playing experiences may be enhanced by correlated haptic, audio, and visual signals when the User Representation hits the ball (Sigrist, Rauter, Riener, & Wolf, 2013). In the future, the use of full-body haptic suits (e.g., Teslasuit[19]), which provide kinesthetic feedback through muscle stimulation or exoskeletons, might provide a more immersive and convincing playing experience. In contrast, input systems that support uninstrumented users, such as the Kinect, do not currently provide an easy opportunity for haptic feedback. This may be remedied in future, for example with the help of unobtrusive wearables, implants, or mid-air haptic feedback (e.g., Ultraleap[20]).

5. Framework and research agenda

Building on the above literature review and examples, we suggest a framework based on theoretical, experimental and application research related to User Representations, and propose a research agenda. We have formulated the present framework and research agenda to cover topics and issues that, from our personal standpoint and a review of the literature, appear to be the most critical and interesting ones. Our goal is to provide an organized vision of User Representations in the context of HCI, which has the potential to yield deeper insights into interaction phenomena. Importantly, in the research agenda, we include relevant questions that have not been fully addressed by past studies, and that can serve as a starting point for future User Representation research and design.

Below we list the theoretical implications of our framework and establish a research agenda for each of the implications discussed.

- **1st Theoretical Implication**: *The visual appearance of a User Representation modifies perception and can enhance or inhibit certain user behaviors, in particular when it elicits illusory body ownership.*

Evidence so far suggests that specific visual features of a User Representation can potentially promote or inhibit certain behaviors. An example applied to desktop environments, is the impact of the directional cues of a cursor on user performance and how users handle a device. In this regard, users tend to move more quickly toward certain locations or to handle an input device differently, depending on the pointing direction of the cursor (Kadri et al., 2007; Po et al., 2005). Recently, several studies related to body ownership and VR have found that the visual appearance of an avatar experienced from 1PP can modulate users' perceptions, attitudes and behaviors. Examples of such phenomena include increased avoidance of a virtual threat when represented by highly realistic hands, compared to iconic hands (Argelaguet et al., 2016), the decrease in implicit racial bias after embodiment of a dark-skinned avatar (Peck et al., 2013), or the execution of more expressive drumming patterns in a VR game when embodied in an avatar that resembles a stereotypical musician, compared to an avatar that looks like a businessman (Kilteni et al., 2013).

Questions for future research:

- How does the visual appearance of a User Representation impact user experience?
- Can the visual appearance of a User Representation be leveraged to foster specific types of behaviors, in order to optimize interaction in specific applications or environments? For instance, in applications which aim to promote physical exercise, User Representations that look strong, motivated, and positive, may potentially enhance physical activity. However, the opposite can also be true, so that these aspects should be further investigated.

---

[19]https://teslasuit.io/.
[20]https://www.ultraleap.com/.

- In what type of applications and contexts can it be beneficial to experience body ownership of a User Representation? When does body ownership of a virtual avatar modulate user behavior? Is this limited to prior user associations regarding certain characteristics of a virtual avatar, or can it evolve during interaction? Does this depend on the purpose of the interaction?
- What are the ethical implications of such implicit modulations of user behaviors and attitudes through the design of a User Representation? Can visual properties be exploited to improve society by putting users virtually in the shoes of other people (e.g., embodying a dark skinned avatar to decrease bias)? What are the potential risks of the technology in this regard (e.g., military purposes, violent games)?
- Can users experience body ownership of User Representations that do not have anthropomorphic characteristics, if they are otherwise plausible and the user is provided with adequate multisensory feedback?
- How do other virtual objects or agents in a virtual environment modulate the perception of a User Representation and its influence on the user's behaviors (e.g., presence of other directional cues, specific behaviors of other avatars interacting with the personal avatar)?

- **2nd Theoretical Implication**: *Visual appearance provides explicit and implicit indications of possible actions and affordances that are associated with a User Representation.*

Virtual User Representations allow quick changes of their semantic properties and functionalities, through adaptation of their appearance. Such representations may entail a simple cursor (e.g., arrow, hand, pencil, brush, eraser, etc.), or a realistic avatar, while potentially supporting the same functionalities. The visual aspects of a User Representation can quickly and effectively convey information to the user about the possible actions and affordances of the representation. This can be done by providing explicit or implicit cues. For example, an explicit cue could be based on using the virtual counterpart of the HTC Vive controllers when immersed in VR, or showing a "scissors" or a "hand" icon, when cutting or selection actions are possible. Explicit cues are thereby related to our everyday use of and associations with objects, e.g., scissors are used to cut, and our hands to grab things. However, it is also possible to leverage implicit associations between the User Representation and the virtual environment that "afford" certain actions. For instance, the size of a virtual hand or tool, can implicitly establish whether the user can pick up a certain object, or the relative size of an avatar pre-defines through which doors the virtual character can pass or the type of obstacles it will be able to overcome.

Questions for future research:

- What are the differences between information conveyed explicitly versus implicitly through a User Representation? Is one method more effective than the other? Does this depend on specific contexts/applications?
- How do associations between a given visual appearance and functionality impact the effectiveness of the interaction? To what extent does the interaction benefit from exploiting associations established through our everyday use of tools and objects (e.g., scissors, erasers, paintbrushes)?
- What are the cognitive mechanisms operating when learning a radically new association between a function and a User Representation? For example, when we learn to use Photoshop magic wand to select pixels based on color, or when we use hands in VR applications like LeapMotion Paint or Gravity Sketch to paint? Are there some visual properties of a representation that can enhance and facilitate the learning process?
- Within a single application, the User Representation can quickly change its shape and semantic properties to suggest certain action possibilities. For instance, a hand-shaped cursor may change to a bi-directional arrow-shape when moving over the scrollbar, or become a vertical line in the text field. Here, the following questions arise: How do such dynamic and quick changes of appearance and

functionality of the representation impact user perceptions? What cognitive skills underlie quick adaptation to constantly changing virtual tools? What would happen if a humanoid avatar's visual appearance were to change quickly to signify new possible actions, analogous to the way a mouse cursor changes?

- **3rd Theoretical Implication**: *Temporal and spatial alignment of multisensory feedback leads to a coherent and meaningful unitary perception during interaction through a User Representation.*

Substantial research has shown that spatial and temporal alignment of different sensory modalities is crucial for the perception of coherent meaningful events. For instance, when users see a virtual object touching their virtual avatar and feel synchronous haptic feedback at the same body location, they have the illusion of being touched by thevirtual object (Bourdin et al., 2017). Thus, the alignment of sensory feedback is crucial for the development of realistic virtual experiences. However, perceptual illusions can also arise as a result of the dominance of one sensory modality over other senses (Chen & Spence, 2017). In this regard, sensory conflicts can be leveraged to make subtle sensory misalignments imperceptible when interacting through a User Representation. For example, the dominance of vision over proprioception can be exploited in VR to induce an illusion of ownership for a virtual body that is not collocated with the user's physical body, in order to reduce fatigue during interaction (Feuchtner & Müller, 2018). Similarly, the dominance of vision can be used to create the illusion of being able to touch and interact with different virtual objects using a virtual hand, although users always touch the same physical prop in the real world (Azmandian et al., 2016; Cheng et al., 2017).

Questions for future research:

- User Representations are virtual objects through which a user executes an action with predicted sensory outcomes (i.e., visual, auditory, tactile). In this context, what is the impact of User Representations on the perception of sensory alignments or misalignments? Can the design of User Representations be optimized to enhance multisensory integration?
- To what extent can sensory mismatches be introduced for specific aims (e.g. reducing fatigue or creating the illusion of being on a roller coaster) without the user noticing them?
- What spatiotemporal characteristics of sensory feedback are more effective when interacting through a User Representation, in order to avoid the perception of mismatches and to create realistic virtual experiences?
- Are sensory misalignments tolerated better when specific characteristics of the User Representation are given, such as a certain visual appearance or method used to control it (i.e., input device, mapping)?

- **4th Theoretical Implication**: *Rich multimodal feedback can impact the user's level of immersion and perceived "naturalness" of the interaction through a User Representation.*

The stimulation of multiple sensory modalities at the same time may enhance the user's feeling of immersion and presence (Slater, 2009), as well as increase the sense of body ownership or tool embodiment regarding specific User Representations. Imagine users playing an immersive tennis game, where, in addition to being able to control a virtual character's motion based on the tracking of body movements, users also hold a real tennis racket in their hands. Moreover, through actuators in the racket, they can feel the impact of balls upon hitting, and even predict the weight of their bodies based on the sound of their steps emitted by shoes with sensors (Tajadura-Jiménez et al., 2015). Rich multimodal feedback allows users to experience the virtual world through their User Representation, using all their senses. Such rich sensory feedback has the potential to lead to faster recalibration of peripersonal space

and tool embodiment. However, future research should also consider how such feedback should be implemented, since recent evidence also shows that if sensory feedback is not sufficiently well-designed and realistic, it can produce feelings of unease, similar to the uncanny valley effect (Berger et al., 2018).

Questions for future research:

- Does stimulating multiple sensory modalities always lead to higher levels of immersion and presence during interaction with a User Representation? Which sensory modalities need to be stimulated, and how, to accomplish this effect?
- How does the richness of the provided sensory feedback (i.e., number of sensory channels stimulated) impact tool-embodiment and body ownership? Can rich multisensory feedback enhance tool embodiment and improve the quality of interaction? For example, is tool embodiment enhanced when an application not only provides visual and auditory feedback, but also includes tactile feedback?
- What are the principles underlying well-designed and realistic sensory feedback? When do feelings of unease toward a User Representation (i.e., uncanny valley) occur? Do these occur in all modalities? How can they be prevented?

- **5th Theoretical Implication**: *The sense of agency over a User Representation is modulated (i.e., enhanced or diminished) by the user's sensory predictions and perceived sensory feedback during interaction.*

Establishing a close match between the sensory feedback provided by an interface and the user's sensory predictions evokes a high sense of agency when interacting via a User Representation (Limerick et al., 2014). We propose that this can be accomplished by designing consistent mappings between user input and User Representation. Noise, latencies, jitter, and a mapping that changes over time can have a negative impact on users' sense of agency (Coyle et al., 2012). However, it should also be noted that in some cases, agency over a User Representation may be intentionally disrupted. For example, this occurs in video games when a level is lost or won: the game might show the avatar celebrating, or dying, there may be a loading sequence, and the player may be transported back to a level selection menu. While such control interruptions do not significantly impact the sense of agency in following gameplay sessions, and may even offer the user a break from "performing" in between levels, providing continuous control without such interruptions might further increase agency.

Another interesting possibility is to enhance agency for a User Representation through training. Tutorials and teaching programs could be developed in order to train user ability to predict the effect of their input on the User Representation. Finally, research has also suggested that some sensory mismatches between predicted and actual sensory outcomes can be tolerated, if they help the user to achieve relevant goals through the interface (e.g., auto-aiming) (Coyle et al., 2012; Kumar & Srinivasan, 2014), or that during body ownership illusions in VR, it is even possible to induce false attributions of agency over an action carried out by another avatar (Banakou & Slater, 2014; Kokkinara et al., 2016). Such aspects challenge traditional theories of agency, and further research should be carried out to better understand why and how it is possible to tolerate given sensory mismatches and to induce agency over actions that "I" did not execute.

Questions for future research:

- What type of techniques (training, tutorials, etc.) can be implemented in order to increase the user's ability to control a User Representation and predict the effect of their actions on the environment?

- What role does the input device and the mappings for user actions in a given interface play for the perceived sense of agency over a User Representation (i.e., keyboard, controller, direct body input)?
- What other factors (e.g., affective, motivational, visual) modulate agency over a User Representation?
- How is it possible to tolerate certain sensory discrepancies between predicted and actual sensory consequences? Does this require users to achieve their high-level goals despite these sensory mismatches?
- Under what circumstances is it possible to let a user experience agency over the actions executed by another avatar, although the user is not actually the agent of such actions (e.g., false attribution of speaking or walking in VR)?

- **6th Theoretical Implication**: *Remapping of the peripersonal space can occur as a consequence of controlling a User Representation to execute actions in virtual environments.*

A remapping of peripersonal space can occur when performing actions in a virtual environment through a User Representation that is visualized in the space external to the user's physical body (Bassolino et al., 2010; Iriki et al., 2001). Such remapping supports the ability to efficiently control virtual objects and to effect changes in the virtual environment (Martel et al., 2016). It should be further investigated how such incorporation of virtual tools into the self-representation can be leveraged to improve interaction. Possibilities include evoking perceptual illusions based on the appearance and other features of the User Representation. For example, distance perception may be modified by interacting through a long virtual arm, or by handling a long virtual tool (Feuchtner & Müller, 2017). Moreover, particular attention should be paid to where in space sensory information is presented. According to reviewed literature, certain sensory information is more efficiently integrated when presented within the peripersonal space (i.e., within a range of approximately 40–70 cm from the operating hands; Holmes & Spence, 2004).

To date, most studies have focused on understanding the modulation of peripersonal space through the use of physical tools, with little research exploring the use of virtual tools (Rallis et al., 2018). In HCI, this has been mainly considered in the context of using a mouse cursor (Bassolino et al., 2010) and recently, when interacting through realistic virtual hands (Alzayat et al., 2019; Bergstrom-Lehtovirta et al., 2019).

Questions for future research:

- Are the mechanisms underlying the remapping of peripersonal space the same when handling physical tools, as when manipulating virtual tools, in particular User Representations? Or are virtual tools, specifically User Representations, incorporated differently into the neural representations of body and space?
- Are all types of User Representation processed in the brain as if they were virtual tools, or is their processing impacted by the type of interfaces and visual appearance of the representation (e.g., human-like, object-like, collocation of the representation, etc.)?
- How can we exploit the incorporation of User Representations into the body schema, for example to change user perceptions of distances or spatial representation in a virtual environment?
- Is the impact of sensory feedback and notification awareness increased when information is presented within the peripersonal space in a virtual environment?
- Does the visual appearance of a User Representation (i.e., body-shaped or tool-shaped virtual representation) impact on how the active control of this virtual tool reshapes the peripersonal space?
- How do 'superpowers' enabled by computers, such as being able to move and control remote objects, impact peripersonal space?
- If the user is working on multiple computer monitors, is it possible that the peripersonal space flexibly remaps, so as to include the area of the screen on which the mouse cursor is being

used? How does the distance of the display from the users affect the quality of the interaction? How can indirect sensory feedback, containing cues related to the configuration of the virtual environment, impact users' spatial perception?

- **7th Theoretical Implication**: *First person perspective (1PP) enhances identification and feelings of connection with a User Representation. When including an HMD and a User Representation with anthropomorphic features, the user can also experience illusory body ownership. A third person perspective (3PP) offers a better situational overview and provides the user with a high degree of spatial awareness of the actions carried out through a User Representation.*

It is important to carefully consider from which visual perspective a User Representation should be experienced, depending on the goals of an interface or application. Based on the presented evidence, we argue that 1PP can increase the user's sense of identification and make them feel strongly connected to their virtual User Representation, whereas 3PP can offer a wider field of view that is advantageous for navigation tasks and spatial awareness. In some applications it can be beneficial to let users switch between visual perspectives as desired (Gorisse et al., 2017). In VR it has been shown that the sense of where the user feels located in space (self-location) can be manipulated by providing specific visuo-tactile and visuo-motor stimulation (Maselli, 2015). This can lead to the perception of being co-located with an avatar that is seen from 3PP (out-of-body illusion). Designers can consider exploiting such illusions to increase the sense of self-identification experienced for a virtual character in 3PP.

Questions for future research:

- When controlling an avatar on a screen, can congruent sensory feedback (e.g., feeling haptic feedback on a body part and perceiving a matching visual stimulus on the avatar's corresponding body part, or making the avatar's body movements match the users' real movements) produce a perceived drift of self-location toward the avatar/screen?
- What are further advantages or disadvantages of experiencing a User Representation from 1PP or 3PP? For example, how does visual perspective affect factors such as cognitive load, performance, emotional state, and engagement?

6. Special cases and limitations

At this point we would like to emphasize that the proposed framework is not exhaustive and does not include all concepts related to User Representations. Nonetheless, it covers a set of topics that we consider relevant for better understanding and designing User Representations in HCI. It should be noted, however, that there are some interesting cases for which it is not straightforward to apply the proposed framework. In this section we describe the limitations of the proposed framework based on some such examples.

Example 1: 3PP Shooter Crosshair

In action games with 3PP, the avatar is the User Representation that players control continuously and through which they can act. However, when aiming a weapon, many games display crosshairs to indicate where the avatar is aiming. The crosshairs then represent the locus of action and the player will focus on controlling the position thereof, sometimes in addition to the avatar's position. The crosshairs could be interpreted as a part or extension of the User Representation. On the other hand, it could also be considered as a secondary User Representation, since it can, to some degree, move independently of the avatar. In cases where the avatar's position is not controlled simultaneously

(i.e., no walking and shooting at the same time), it could be argued that the avatar ceases to be the User Representation, and that the crosshairs take over this role.

Example 2: Control of the World instead of the Avatar

Some games, such as Super Monkey Ball,[21] feature a virtual avatar in the center of the screen (in this example a monkey in a ball), which cannot be controlled directly. Instead, its movements are controlled by tilting the platforms it is on, making the ball roll down the slope. Such controls are frequently mapped to a corresponding tilting of the input device. The main focus of players is the ball at the center of the screen, which they try to navigate along a path to reach a target. The ball represents the locus of action, in that collectibles (bananas, coins, etc.) are picked up when it touches them. The game is won when it reaches the goal, and lost when it falls off the platforms. While the ball movements can be controlled continuously, they are controlled indirectly. According to our definition, the ball qualifies as the User Representation. Nevertheless, it could also be argued that the platforms constitute the User Representation, since they are in fact controlled by the player and therefore enable actions in the virtual environment.

Example 3: Text Cursor

The text cursor, or caret, is the vertical, flashing line between text characters, which indicates the position at which text can be entered or removed. While typing, this caret automatically moves ahead of the next typed character and can otherwise be moved with the arrow keys (caret navigation). The caret may also be positioned by clicking into a line of text or a text field with the mouse cursor, whereby the mouse cursor usually also has a caret shape. Commonly the mouse cursor and the text cursor are simultaneously visible, and since both fulfill the purpose of signifying the user's location for action, they could both be considered User Representations. However, it could also be argued that the text cursor is not controlled continuously, since it moves to new positions through discrete key presses. Thus the mouse cursor would remain the single User Representation. On the other hand, while the user is typing, the mouse cursor is no longer the locus of interaction and the center of focus, as a result of which it might temporarily cease to be the User Representation.

Example 4: Selection of text or graphical elements in 2D screen space

In the example of the text editing application, the user can also select consecutive characters, words, lines, or even pages, by clicking and dragging. Consequently, the selected text is marked by a colored highlight, and following operations will apply to the full selection (e.g., change of text style, copy, cut, delete). Similarly, area cursors (Kabbash & Buxton, 1995) are often used in techniques like rectangle or lasso selection, for selecting multiple graphical elements on a 2D screen space (e.g., Photoshop). Despite the selected text indicating the user's current focus and point of interaction, it cannot be considered a User Representation based on the presented definition. The reason is that the selection is only a temporary representation of the user's current target of manipulation, not however, a continuous proxy for interaction. Instead, the cursor that is used to create this selection remains the User Representation, even if it may temporarily be invisible.

Example 5: User Representation on a Touch Screen

Touch screens typically do not provide virtual User Representations, since the users directly manipulate the target of interest with their fingers or a stylus. Due to the resulting lack of virtual User Representations, touch screen interaction is not discussed in depth in this paper. An exception worth mentioning is the work by Costes, Argelaguet, Danieau, Guillotel, and Lécuyer (2019), which presents an approach for simulating haptic feedback through a virtual User Representation on a touch screen. The user is represented by a circular cursor at the touch

---

[21]https://games.sega.com/supermonkeyball/.

point, which dynamically changes in shape and size to reflect different material properties. For example, when swiping over a rough surface, the cursor will slightly jitter up and down; to signal pressing down on soft surfaces, the radius of the cursor changes over time; and slippery surfaces are indicated by velocity decoupling between finger and cursor.

With these examples, we argue that our definition of User Representations is not capable of covering all forms of computer interactions. Instead we aim to provide a framework for categorization, analysis and design of common computer interactions, which may be built upon and extended to include exceptions, and can serve as a basis for the discussion thereof.

7. Conclusion

In this paper we propose the concept of User Representations to analyze user perceptions of cursors, avatars, virtual hands, or tools and similar virtual objects from a common perspective. User Representations are virtual objects that serve as artificial extensions of the users' physical bodies, enabling them to execute actions in virtual environments, which would otherwise be unreachable. Users can continuously control these User Representations through their motor commands, leading to a perceived sense of agency over the representation. These virtual objects can vary in several aspects, such as experienced visual perspective, appearance, and the sensory feedback provided. Further, the User Representation may be controlled through different input devices and mapping functions, which can affect the experienced sense of agency, remapping of peripersonal space, and in some cases the illusion of body ownership. In this context, we have shown how knowledge from psychology, neuroscience, and HCI can help us understand how users perceive their virtual User Representations. We are confident that our work will help researchers in HCI gain a deeper understanding of interaction through artificial user representations, as well as inspire future research on these topics. To this end, we propose a research agenda for these concepts that can provide the HCI community with a wider perspective on how the interaction of a user with a computer can be modulated by the design of their User Representations.

Notes on contributors

Sofia Seinfeld (sofia.seinfeld@uni-bayreuth.de; LinkedIn Profile: Twitter) is a psychologist with an expertise in perception, multisensory integration, and virtual reality. Currently, she is a postdoctoral researcher in Computer Science at the University of Bayreuth, with an interest in HCI and the exploitation of multisensory integration principles to improve user experience and interface design.

Tiare Feuchtner (tiare@cs.au.dk; LinkedIn Profile; Twitter) is a computer scientist with main interest in HCI, in particular interaction in immersive virtual environments. She is currently a postdoctoral researcher at the Department of Computer Science of Aarhus University.

Antonella Maselli (antonella.maselli@istc.cnr.it; LinkedIn Profile) is a neuroscientist with interests in body illusions, multisensory integration, motion perception, motor control and non-verbal interpersonal communication. She is currently a researcher at the Institute of Cognitive Sciences and Technologies at the CNR in Rome, Italy.

Jörg Müller (joerg.mueller@uni-bayreuth.de, http://joergmueller.info) is a computer scientist with an interest in HCI. He is a Professor of Computer Science at the University of Bayreuth, Germany.

ORCID

Sofia Seinfeld ⓘ http://orcid.org/0000-0001-9649-0785

References

Achibet, M., Girard, A., Talvas, A., Marchal, M., & Lecuyer, A. (2015). *Elastic-Arm: Human-scale passive haptic feedback for augmenting interaction and perception in virtual environments. Proceedings of the IEEE VR 2015 Conference on Virtual Reality and 3D User Interfaces*. New York: IEEE Computer Society. doi: 10.1109/VR.2015.7223325

Ahn, S. J. G., Bostick, J., Ogle, E., Nowak, K. L., McGillicuddy, K. T., & Bailenson, J. N. (2016). Experiencing nature: Embodying animals in immersive virtual environments increases inclusion of nature in self and involvement with nature. *Journal of Computer-Mediated Communication*, 21(6), 399–419. doi:10.1111/jcc4.12173

Akamatsu, M., Sato, S., & MacKenzie, I. S. (1994). Multimodal mouse: A mouse-type device with tactile and force display. *Presence: Teleoperators and Virtual Environments*, 3(1), 73–80. doi:10.1162/pres.1994.3.1.73

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14 (3), 257–262. doi:10.1016/J.CUB.2004.01.029

Alzayat, A., Hancock, M., & Nacenta, M. A. (2019). *Quantitative measurement of tool embodiment for virtual reality input alternatives. Proceedings of the CHI 2019 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/3290605.3300673

Arbib, M. A., Bonaiuto, J. B., Jacobs, S., & Frey, S. H. (2009). Tool use and the distalization of the end-effector. *Psychological Research Psychologische Forschung*, 73(4), 441–462. doi:10.1007/s00426-009-0242-2

Argelaguet, F., Hoyet, L., Trico, M., & Lecuyer, A. (2016). *The role of interaction in virtual embodiment: Effects of the virtual hand representation. Proceedings of the IEEE VR 2016 Virtual Reality Conference*. New York: IEEE Computer Society. doi: 10.1109/VR.2016.7504682

Asai, T. (2015). Feedback control of one's own action: Self-other sensory attribution in motor control. *Consciousness and Cognition*, 38, 118–129. doi:10.1016/j.concog.2015.11.002

Ash, E. (2016). Priming or proteus effect? Examining the effects of avatar race on in-game behavior and post-play aggressive cognition and affect in video games. *Games and Culture*, 11(4), 422–440. doi:10.1177/1555412014568870

Azmandian, M., Hancock, M., Benko, H., Ofek, E., & Wilson, A. D. (2016). *Haptic retargeting. Proceedings of the CHI 2016 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/2858036.2858226

Bachynskyi, M., Paneva, V., & Müller, J. (2018). *LeviCursor. Proceedings of the ISS 2018 International Conference on Interactive Surfaces and Spaces*. New York: ACM. doi:10.1145/3279778.3279802

Baecker, R., & Buxton, W. (1987, January 1). *Readings in human-computer interaction: A multidisciplinary approach*. Los Altos, CA: Morgan Kaufman Publishers, Inc.. Retrieved from https://www.osti.gov/biblio/5039704

Bærentsen, K. B., & Trettvik, J. (2002). *An activity theory approach to affordance. Proceedings of the NordiCHI 2002 Nordic Conference on Human-Computer Interaction*. New York: ACM. doi: 10.1145/572020.572028

Balakrishnan, R. (2004). "Beating" Fitts' law: Virtual enhancements for pointing facilitation. *International Journal of Human-computer Studies*, 61(6), 857–874. doi:10.1016/J.IJHCS.2004.09.002

Ban, Y., Narumi, T., Tanikawa, T., & Hirose, M. (2013). Modifying perceived size of a handled object through hand image deformation. *Presence: Teleoperators and Virtual Environments*, 22(3), 255–270. doi:10.1162/PRES_a_00154

Banakou, D., Groten, R., & Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, 110(31), 12846–12851. doi:10.1073/pnas.1306779110

Banakou, D., Kishore, S., & Slater, M. (2018). Virtually being Einstein results in an improvement in cognitive task performance and a decrease in age bias. *Frontiers in Psychology*, 9. doi:10.3389/fpsyg.2018.00917

Banakou, D., & Slater, M. (2014). Body ownership causes illusory self-attribution of speaking and influences subsequent real speaking. *Proceedings of the National Academy of Sciences*, 111(49), 17678–17683. doi:10.1073/pnas.1414936111

Bashford, L., & Mehring, C. (2016). Ownership and agency of an independent supernumerary hand induced by an imitation brain-computer interface. *PloS One*, 11(6), e0156591. doi:10.1371/journal.pone.0156591

Bassolino, M., Serino, A., Ubaldi, S., & Làdavas, E. (2010). Everyday use of the computer mouse extends peripersonal space representation. *Neuropsychologia*, 48(3), 803–811. doi:10.1016/J.NEUROPSYCHOLOGIA.2009.11.009

Beaudouin-Lafon, M. (2000). *Instrumental interaction. Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/332040.332473

Berger, C. C., Gonzalez-Franco, M., Ofek, E., & Hinckley, K. (2018). The uncanny valley of haptics. *Science Robotics*, 3 (17), eaar7010. doi:10.1126/scirobotics.aar7010

Bergstrom-Lehtovirta, J., Mottelson, A., Muresan, A., & Hornbæk, K. (2019). *Tool extension in human-computer interaction.* Proceedings of the CHI 2004 Conference on Human Factors in Computing Systems. New York: ACM. doi: 10.1145/3290605.3300798

Berti, A., & Frassinetti, F. (2000). When far becomes near: remapping of space by tool use. *Journal of Cognitive Neuroscience*, 12(3), 415–420. doi:10.1162/089892900562237

Blanch, R., Guiard, Y., & Beaudouin-Lafon, M. (2004). *Semantic pointing. Proceedings of the CHI 2004 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/985692.985758

Blanke, O., & Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences*, 13(1), 7–13. doi:10.1016/j.tics.2008.10.003

Bødker, S. (1987). Through the interface - a human activity approach to user interface design. *DAIMI Report Series*, 16 (224). doi:10.7146/dpb.v16i224.7586. Retrieved from https://tidsskrift.dk/daimipb/article/view/7586

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, 391(6669), 756. doi:10.1038/35784

Bourdin, P., Barberia, I., Oliva, R., Slater, M., & Passingham, R. (2017). A virtual out-of-body experience reduces fear of death. *PloS One*, 12(1), e0169343. doi:10.1371/journal.pone.0169343

Braun, N., Debener, S., Spychala, N., Bongartz, E., Sörös, P., Müller, H. H. O., & Philipsen, A. (2018). The senses of agency and ownership: A review. *Frontiers in Psychology*, 9, 535. doi:10.3389/fpsyg.2018.00535

Burigat, S., Chittaro, L., & Sioni, R. (2017). Mobile three-dimensional maps for wayfinding in large and complex buildings: Empirical comparison of first-person versus third-person perspective. *IEEE Transactions on Human-Machine Systems*, 47(6), 1029–1039. doi:10.1109/THMS.2017.2693684

Burr, D., & Alais, D. (2006). Chapter 14 combining visual and auditory information. *Progress in Brain Research*, 155, 243–258. doi:10.1016/S0079-6123(06)55014-9

Calvert, G. A., Spence, C., & Barry, S. E. (2004). *The handbook of multisensory processes*. Massachussets: Massachussetts Institute of Technology.

Card, S. K., Mackinlay, J. D., & Robertson, G. G. (1990). *The design space of input devices. Proceedings of the CHI 1990 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/97243.97263

Cardinali, L., Frassinetti, F., Brozzoli, C., Urquizar, C., Roy, A. C., & Farnè, A. (2009). Tool-use induces morphological updating of the body schema. *Current Biology*, 19(12), R478–R479. doi:10.1016/J.CUB.2009.05.009

Casiez, G., Vogel, D., Balakrishnan, R., & Cockburn, A. (2008). The impact of control-display gain on user performance in pointing tasks. *Human-Computer Interaction*, 23(3), 215–250. doi:10.1080/07370020802278163

Chen, Y. C., & Spence, C. (2017). Assessing the role of the 'Unity Assumption' on multisensory integration: A review. *Frontiers in Psychology*, 8, 445. doi:10.3389/fpsyg.2017.00445

Cheng, L.-P., Ofek, E., Holz, C., Benko, H., & Wilson, A. D. (2017). *Sparse Haptic proxy. Proceedings of the CHI 2017 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/3025453.3025753

Cornelio-Martinez, P., De Pirro, S., Vi, C. T., & Subramanian, S. (2017). *Agency in Mid-air Interfaces. Proceedings of the CHI 2017 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/3025453.3025457

Cornelio-Martinez, P., Maggioni, E., Hornbæk, K., Obrist, M., & Subramanian, S. (2018). *Beyond the libet clock: Modality variants for agency measurements. Proceedings of the CHI 2018 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/3173574.3174115

Costes, A., Argelaguet, F., Danieau, F., Guillotel, P., & Lécuyer, A. (2019). Touchy : A visual approach for simulating haptic effects on touchscreens. *Frontiers in ICT*, 6, 1. doi:10.3389/fict.2019.00001

Coyle, D., Moore, J., Kristensson, P. O., Fletcher, P., & Blackwell, A. (2012). *I did that! Measuring users' Experience of agency in their own actions. Proceedings of the CHI 2012 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1094/PDIS-11-11-0999-PDN

De Souza, C. S. (2005). Semiotic engineering: Bringing designers and users together at interaction time. *Interacting with Computers*, 17(3), 317–341. doi:10.1016/j.intcom.2005.01.007

Denisova, A., & Cairns, P. (2015). *First Person vs. Third Person perspective in digital games. Proceedings of the CHI 2015 Conference on Human Factors in Computing Systems*. New York: ACM. doi:10.1145/2702123.2702256

Di Pellegrino, G., & Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia*, 66, 126–133. doi:10.1016/J.NEUROPSYCHOLOGIA.2014.11.011

Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science (New York, N.Y.)*, 317(5841), 1048. doi:10.1126/science.1142175

Ehrsson, H. H., Spence, C., & Passingham, R. E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science (New York, N.Y.)*, 305(5685), 875–877. doi:10.1126/science.1097011

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8(4), 162–169. doi:10.1016/J.TICS.2004.02.002

Farnè, A., & Làdavas, E. (2002). Auditory peripersonal space in humans. *Journal of Cognitive Neuroscience*, 14(7), 1030–1043. doi:10.1162/089892902320474481

Fernandez, R. A. S., Sanchez-Lopez, J. L., Sampedro, C., Bavle, H., Molina, M., & Campoy, P. (2016). *Natural user interfaces for human-drone multi-modal interaction. Proceedings of ICUAS* 2016 International Conference on Unmanned Aircraft Systems. New York: IEEE Computer Society. doi:10.1109/ICUAS.2016.7502665

Ferri, F., Chiarelli, A. M., Merla, A., Gallese, V., & Costantini, M. (2013). The body beyond the body: Expectation of a sensory event is enough to induce ownership over a fake hand. *Proceedings. Biological Sciences*, *280*(1765), 20131140. doi:10.1098/rspb.2013.1140

Feuchtner, T., & Müller, J. (2017). *Extending the body for interaction with reality. Proceedings of the CHI 2017 Conference on Human Factors in Computing Systems*. New York: ACM. doi: 10.1145/3025453.3025689

Feuchtner, T., & Müller, J. (2018). *Ownershift: Facilitating overhead interaction in virtual reality with an ownership-preserving hand space shift. Proceedings of the UIST 2018 Symposium on User Interface Software and Technology*, New York: ACM. doi:10.1145/3242587.3242594

Folegatti, A., de Vignemont, F., Pavani, F., Rossetti, Y., & Farnè, A. (2009). Losing One's Hand: Visual-proprioceptive conflict affects touch perception. *PloS One*, *4*(9), e6920. doi:10.1371/journal.pone.0006920

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1404), 1771–1788. doi:10.1098/rstb.2000.0734

Galvan Debarba, H., Bovet, S., Salomon, R., Blanke, O., Herbelin, B., & Boulic, R. (2017). Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PloS One*, *12*(12), e0190109. doi:10.1371/journal.pone.0190109

Georgiou, O., Jeffrey, C., Chen, Z., Xiao Tong, B., Hei Chan, S., Yang, B., … Carter, T. (2018). *Touchless haptic feedback for VR rhythm games. Proceedings of the IEEE VR 2018 Conference on Virtual Reality and 3D User Interfaces*. New York: IEEE Computer Society. doi: 10.1109/VR.2018.8446619

Gibson, J. (1966). *The senses considered as perceptual systems*. Oxford, England: Houghton Mifflin. Retrieved from https://psycnet.apa.org/record/1966-35026-000

Gibson, J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing*. (pp. 127–143). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibson, J. (1979). *The ecological approach to visual perception*. Psychology Press. doi:10.4324/9781315740218

González-Franco, M., Peck, T. C., Rodríguez-Fornells, A., & Slater, M. (2014). A threat to a virtual hand elicits motor cortex activation. *Experimental Brain Research*, *232*(3), 875–887. doi:10.1007/s00221-013-3800-1

Gonzalez-Franco, M., Perez-Marcos, D., Spanlang, B., & Slater, M. (2010). *The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. Proceedings of the IEEE VR 2010 Virtual Reality Conference*. New York: IEEE Computer Society. doi:10.1109/VR.2010.5444805

Gorisse, G., Christmann, O., Amato, E. A., & Richir, S. (2017). First- and Third-Person perspectives in immersive virtual environments: presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, *4*, 33. doi:10.3389/frobt.2017.00033

Gozli, D. G., & Brown, L. E. (2011). Agency and control for the integration of a virtual tool into the peripersonal space. *Perception*, *40*(11), 1309–1319. doi:10.1068/p7027

Graziano, M. S., Yap, G. S., & Gross, C. G. (1994). Coding of visual space by premotor neurons. *Science (New York, N. Y.) 266*(5187): 1054–1057. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7973661

Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging*, *23*(4), 692–701. doi:10.1037/a0014345

Grubert, J., Witzani, L., Ofek, E., Pahud, M., Kranz, M., & Kristensson, P. O. (2018). *Effects of hand representations for typing in virtual reality. Proceedings of the IEEE VR 2018 Conference on Virtual Reality and 3D User Interfaces*. New York: IEEE Computer Society. doi:10.1109/VR.2018.8446250

Guterstam, A., Björnsdotter, M., Gentile, G., & Ehrsson, H. H. (2015). Posterior cingulate cortex integrates the senses of self-location and body ownership. *Current Biology*, *25*(11), 1416–1425. doi:10.1016/J.CUB.2015.03.059

Guterstam, A., Petkova, V. I., & Ehrsson, H. H. (2011). The illusion of owning a third arm. *PloS One*, *6*(2), e17208. doi:10.1371/journal.pone.0017208

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196–207. doi:10.1038/nrn.2017.14

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385. doi:10.1038/nn827

Hirao, Y., & Kawai, T. (2018). Augmented cross-modality: Translating the physiological responses, knowledge and impression to audio-visual information in virtual reality. *Journal of Imaging Science and Technology*, *62*, 6. doi:10.2352/J.ImagingSci.Technol.2018.62.6.060402

Hirayama, R., Martinez Plasencia, D., Masuda, N., & Subramanian, S. (2019). A volumetric display for visual, tactile and audio presentation using acoustic trapping. *Nature*, *575*(7782), 320–323. doi:10.1038/s41586-019-1739-5

Ho, C., & Spence, C. (2009). Using peripersonal warning signals to orient a driver's gaze. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *51*(4), 539–556. doi:10.1177/0018720809341735

Holmes, N. P., Calvert, G. A., & Spence, C. (2004). Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neuroscience Letters*, *372*(1–2), 62–67. doi:10.1016/J.NEULET.2004.09.024

Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation(s) of peripersonal space. *Cognitive Processing*, *5*(2), 94–105. doi:10.1007/s10339-004-0013-3

Hornbæk, K., & Oulasvirta, A. (2017). *What is interaction? Proceedings of the CHI 2017 Conference on Human Factors in Computing Systems.* New York: ACM. doi:10.1145/3025453.3025765

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human–Computer Interaction*, *1*(4), 311–338. doi:10.1207/s15327051hci0104_2

Hwang, I., Son, H., & Kim, J. R. (2017). *AirPiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback. Proceedings of the IEEE WHC 2017 World Haptics Conference .* New York: IEEE Computer Society. doi:10.1109/WHC.2017.7989903

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., … Blanke, O. (2011). Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron*, *70*(2), 363–374. doi:10.1016/J.NEURON.2011.03.009

Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport 7*(14): 2325–2330. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8951846

Iriki, A., Tanaka, M., Obayashi, S., & Iwamura, Y. (2001). Self-images in the video monitor coded by monkey intraparietal neurons. *Neuroscience Research*, *40*(2), 163–173. doi:10.1016/S0168-0102(01)00225-5

Jung, S., Bruder, G., Wisniewski, P. J., Sandor, C., & Hughes, C. E. (2018). Over my hand: Using a personalized hand in VR to improve object size estimation, body ownership, and presence. *Proceedings of the SUI 2018 Symposium on Spatial User Interaction*. New York: ACM. doi: 10.1145/3267782.3267920

Kabbash, P., & Buxton, W. A. S. (1995). *The "prince" technique.* (pp. 273–279). Association for Computing Machinery (ACM). doi:10.1145/223904.223939

Kadri, A., Lecuyer, A., Burkhardt, J., & Richir, S. (2007). *The visual appearance of user's Avatar Can influence the manipulation of both real devices and virtual objects. Proceedings of the 3DUI 2007 Symposium on 3D User Interfaces.* New York: IEEE Computer Society. doi:10.1109/3DUI.2007.340767

Kalckert, A., & Ehrsson, H. H. (2014). The moving rubber hand illusion revisited: Comparing movements and visuotactile stimulation to induce illusory ownership. *Consciousness and Cognition*, *26*, 117–132. doi:10.1016/J.CONCOG.2014.02.003

Kaptelinin, V., & Nardi, B. (2012). *Affordances in HCI. Proceedings of the CHI 2012 Conference on Human Factors in Computing Systems.* New York: ACM. doi: 10.1145/2207676.2208541

Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., & Tachibana, K. (2000). *Virtual object manipulation on a table-top AR environment. Proceedings of the ISAR 2000 International Symposium on Augmented Reality.* New York: IEEE Computer Society. doi:10.1109/ISAR.2000.880934

Kilteni, K., Bergstrom, I., & Slater, M. (2013). Drumming in immersive virtual reality: The body shapes the way we play. *IEEE Transactions on Visualization and Computer Graphics*, *19*(4), 597–605. doi:10.1109/TVCG.2013.29

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373–387. doi:10.1162/PRES_a_00124

Kilteni, K., Maselli, A., Kording, K. P., & Slater, M. (2015). Over my fake body: Body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in Human Neuroscience*, *9*, 141. doi:10.3389/fnhum.2015.00141

Kilteni, K., Normand, J.-M., Sanchez-Vives, M. V., & Slater, M. (2012). Extending body space in immersive virtual reality: A very long arm illusion. *PloS One*, *7*(7), e40867. doi:10.1371/journal.pone.0040867

Kishore, S., Navarro, X., Dominguez, E., De La Peña, N., & Slater, M. (2018). Beaming into the news: A system for and case study of tele-immersive journalism. *IEEE Computer Graphics and Applications*, *38*(2), 89–101. doi:10.1109/MCG.2017.2801407

Klokmose, C. N., & Beaudouin-Lafon, M. (2009). *VIGO: Instrumental interaction in multi-surface environments. Proceedings of the CHI 2009 Conference on Human Factors in Computing Systems.* New York: ACM. doi:10.1145/1518701.1518833

Kokkinara, E., Kilteni, K., Blom, K. J., & Slater, M. (2016). First person perspective of seated participants over a walking virtual body leads to illusory agency over the walking. *Scientific Reports*, *6*(1), 28879. doi:10.1038/srep28879

Kokkinara, E., & Slater, M. (2014). Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception*, *43*(1), 43–58. doi:10.1068/P7545

Kondo, R., Sugimoto, M., Minamizawa, K., Hoshi, T., Inami, M., & Kitazaki, M. (2018). Illusory body ownership of an invisible body interpolated between virtual hands and feet via visual-motor synchronicity. *Scientific Reports*, *8*(1), 7541. doi:10.1038/s41598-018-25951-2

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247. doi:10.1038/nature02169

Kumar, D., & Srinivasan, N. (2014). Naturalizing sense of agency with a hierarchical event-control approach. *PloS One*, *9*(3), e92431. doi:10.1371/journal.pone.0092431

Lécuyer, A. 2009, February. Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence: Teleoperators and Virtual Environments*, *18*, 39–53. doi:10.1162/pres.18.1.39

Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: Manipulating bodily self-consciousness. *Science (New York, N.Y.)*, *317*(5841), 1096–1099. doi:10.1126/science.1143439

Liepelt, R., Dolk, T., & Hommel, B. (2017). Self-perception beyond the body: The role of past agency. *Psychological Research*, *81*(3), 549–559. doi:10.1007/s00426-016-0766-1

Limerick, H., Coyle, D., & Moore, J. W. (2014). The experience of agency in human-computer interactions: A review. *Frontiers in Human Neuroscience*, *8*, 643. doi:10.3389/fnhum.2014.00643

Limerick, H., Moore, J. W., & Coyle, D. (2015). *Empirical evidence for a diminished sense of agency in speech interfaces*. Proceedings of the CHI 2015 Conference on Human Factors in Computing Systems. New York: ACM. doi: 10.1145/2702123.2702379

Lin, L., & Jorg, S. (2016). *The effect of realism on the virtual hand illusion. Proceedings of the IEEE VR 2016 Conference on Virtual Reality and 3D User Interfaces*. New York: IEEE Computer Society. doi: 10.1109/VR.2016.7504731

Lin, L., Normovle, A., Adkins, A., Sun, Y., Robb, A., Ye, Y., … Jörg, S. (2019). *The effect of hand size and interaction modality on the virtual hand illusion. Proceedings of the IEEE VR 2019 Conference on Virtual Reality and 3D User Interfaces*. New York: IEEE Computer Society. doi: 10.1109/VR.2019.8797787

Linkenauger, S. A., Leyrer, M., Bülthoff, H. H., & Mohler, B. J. (2013). Welcome to Wonderland: The influence of the size and shape of a virtual hand on the perceived size and shape of virtual objects. *PloS One*, *8*(7), e68594. doi:10.1371/journal.pone.0068594

Ma, K., & Hommel, B. (2015). Body-ownership for actively operated non-corporeal objects. *Consciousness and Cognition*, *36*, 75–86. doi:10.1016/J.CONCOG.2015.06.003

Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2014). Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences*, *19*(1), 6–12. doi:10.1016/j.tics.2014.11.001

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, *8*(2), 79–86. doi:10.1016/J.TICS.2003.12.008

Marshall, J., Benford, S., Byrne, R., & Tennent, P. (2019). *Sensory alignment in immersive entertainment. Proceedings of the CHI 2019 Conference on Human Factors in Computing Systems*. New York: ACM. doi: 10.1145/3290605.3300930

Martel, M., Cardinali, L., Roy, A. C., & Farnè, A. (2016). Tool-use: An open window into body representation and its plasticity. *Cognitive Neuropsychology*, *33*(1–2), 82–101. doi:10.1080/02643294.2016.1167678

Maselli, A. (2015). Allocentric and egocentric manipulations of the sense of self-location in full-body illusions and their relation with the sense of body ownership. *Cognitive Processing*, *16*(S1), 309–312. doi:10.1007/s10339-015-0667-z

Maselli, A., Kilteni, K., López-Moliner, J., & Slater, M. (2016). The sense of body ownership relaxes temporal constraints for multisensory integration. *Scientific Reports*, *6*(1), 30628. doi:10.1038/srep30628

Maselli, A., & Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience*, *7*, 83. doi:10.3389/fnhum.2013.00083

Maselli, A., & Slater, M. (2014). Sliding perspectives: Dissociating ownership from self-location during full body illusions in virtual reality. *Frontiers in Human Neuroscience*, *8*, 693. doi:10.3389/fnhum.2014.00693

Matamala-Gomez, M., Diaz Gonzalez, A. M., Slater, M., & Sanchez-Vives, M. V. (2018). Decreasing pain ratings in chronic arm pain through changing a virtual body: Different strategies for different pain types. *The Journal of Pain*. doi:10.1016/J.JPAIN.2018.12.001

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662. doi:10.1152/jn.1986.56.3.640

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*(8), 1265–1279. doi:10.1016/S0893-6080(96)00035-4

Michotte, A. (2017). *The perception of causality*. New York: Basic Books. doi:10.4324/9781315519050

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*(1), 546–561. doi:10.1016/j.concog.2011.12.002

Müller, J., Oulasvirta, A., & Murray-Smith, R. (2017). Control theoretic models of pointing. *ACM Transactions on Computer-Human Interaction*, *24*(4), 1–36. doi:10.1145/3121431

Müller, J., Walter, R., Bailly, G., Nischt, M., & Alt, F. (2012). *Looking glass. Proceedings of the CHI 2012 Conference on Human Factors in Computing Systems*. New York: ACM. doi: 10.1145/2207676.2207718

Nierula, B., Martini, M., Matamala-Gomez, M., Slater, M., & Sanchez-Vives, M. V. (2017). Seeing an embodied virtual hand is analgesic contingent on colocation. *The Journal of Pain*, *18*(6), 645–655. doi:10.1016/J.JPAIN.2017.01.003

Noel, J.-P., Samad, M., Doxon, A., Clark, J., Keller, S., & Di Luca, M. (2018). Peri-personal space as a prior in coupling visual and proprioceptive signals. *Scientific Reports*, *8*(1), 15819. doi:10.1038/s41598-018-33961-3

Norman, D. (1988). *The psychology of everyday things*. New York, NY: Doubleday. Retrieved from http://www.aisacademics.com/wp-content/uploads/2016/07/IND133-Human-Factors.pdf

Norman, D. (1999). Affordance, conventions, and design. *Interactions*, *6*(3), 38–43. doi:10.1145/301153.301168

Norman, D. (2008). The way I see it: Signifiers, not affordances. *Interactions*, *15*(6), 18. doi:10.1145/1409040.1409044

Obrist, M., Velasco, C., Vi, C., Ranasinghe, N., Israr, A., Cheok, A., … Gopalakrishnakone, P. (2016). Sensing the future of HCI. *Interactions*, *23*(5), 40–49. doi:10.1145/2973568

Ogawa, N., Narumi, T., & Hirose, M. (2019). *Virtual hand realism affects object size perception in body-based scaling. Proceedings of the IEEE VR 2019 Conference on Virtual Reality and 3D User Interfaces*. New York: IEEE Computer Society. Doi: 10.1109/VR.2019.8798040

Oyanagi, A., & Ohmura, R. (2017). Investigating the sense of body ownership over a bird avatar for enhancing immersion in flying experience. *Transactions of the Virtual Reality Society of Japan*, *22*(4), 513–522. doi:10.18974/tvrsj.22.4_513

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787. doi:10.1016/j.concog.2013.04.016

Petkova, V. I., & Ehrsson, H. H. (2008). If I were you: perceptual illusion of body swapping. *PloS One*, *3*(12), e3832. doi:10.1371/journal.pone.0003832

Picard, F., & Friston, K. (2014). Predictions, perception, and a sense of self. *Neurology*, *83*(12), 1112–1118. doi:10.1212/WNL.0000000000000798

Po, B. A., Fisher, B. D., & Booth, K. S. (2005). *Comparing cursor orientations for mouse, pointer, and pen interaction. Proceedings of the CHI 2005 Conference on Human Factors in Computing Systems*. New York: ACM. doi: 10.1145/1054972.1055013

Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). *The go-go interaction technique. Proceedings of the UIST 1996 Symposium on User Interface Software and Technology*. New York: ACM. doi: 10.1145/237091.237102

Previc, F. H. (2000). Neuropsychological guidelines for aircraft control stations. *IEEE Engineering in Medicine and Biology Magazine*, *19*(2), 81–88. doi:10.1109/51.827410

Rallis, A., Fercho, K. A., Bosch, T. J., & Baugh, L. A. (2018). Getting a handle on virtual tools: An examination of the neuronal activity associated with virtual tool use. *Neuropsychologia*, *109*, 208–221. doi:10.1016/J.NEUROPSYCHOLOGIA.2017.12.023

Salisbury, J. K., & Srinivasan, M. A. (1997). Phantom-based haptic interaction with virtual objects. *IEEE Computer Graphics and Applications*, *17*(5), 6–10. doi:10.1109/MCG.1997.1626171

Samad, M., Gatti, E., Hermes, A., Benko, H., & Parise, C. (2019). *Pseudo-haptic weight: Changing the perceived weight of virtual objects by manipulating control-display ratio. Proceedings of the CHI 2019 Conference on Human Factors in Computing Systems*. New York: ACM. doi: 10.1145/3290605.3300550

Sanchez-Vives, M. V., Spanlang, B., Frisoli, A., Bergamasco, M., & Slater, M. (2010). Virtual hand illusion induced by visuomotor correlations. *PloS One*, *5*(4), e10381. doi:10.1371/journal.pone.0010381

Schwind, V., Lin, L., Di Luca, M., Jörg, S., & Hillis, J. (2018). *Touch with foreign hands: The effect of virtual hand appearance on visual-haptic integration. Proceedings of the SAP 2018 Symposium on Applied Perception*. New York: ACM. doi: 10.1145/3225153.3225158

Seinfeld, S., Arroyo-Palacios, J., Iruretagoyena, G., Hortensius, R., Zapata, L. E., Borland, D., … Sanchez-Vives, M. V. (2018). Offenders become the victim in virtual reality: Impact of changing perspective in domestic violence. *Scientific Reports*, *8*(1), 2692. doi:10.1038/s41598-018-19987-7

Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., & Lopez, C. (2013). Bodily ownership and self-location: Components of bodily self-consciousness. *Consciousness and Cognition*, *22*(4), 1239–1252. doi:10.1016/J.CONCOG.2013.08.013

Serino, A., Noel, J.-P., Mange, R., Canzoneri, E., Pellencin, E., Ruiz, J. B., … Herbelin, B. (2018). Peripersonal space: An index of multisensory body–environment interactions in real, virtual, and mixed realities. *Frontiers in ICT*, *4*, 31. doi:10.3389/fict.2017.00031

Shneiderman, B. (1997). *Direct manipulation for comprehensible, predictable and controllable user interfaces. Proceedings of the IUI 97 Conference on Intelligent User Interfaces*. New York: ACM. doi: 10.1145/238218.238281

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic Bulletin and Review*. doi:10.3758/s13423-012-0333-8

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557. doi:10.1098/rstb.2009.0138

Slater, M., Pérez Marcos, D., Ehrsson, H., & Sanchez-Vives, M. V. (2009). Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience*, *3*(2), 214–220. doi:10.3389/neuro.01.029.2009

Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). First person experience of body transfer in virtual reality. *PloS One*, *5*(5), e10564. doi:10.1371/journal.pone.0010564

Swapp, D., Pawar, V., & Loscos, C. (2006). Interaction with co-located haptic feedback in virtual reality. *Virtual Reality*, *10*(1), 24–30. doi:10.1007/s10055-006-0027-5

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219–239. doi:10.1016/J.CONCOG.2007.03.010

Taima, Y., Ban, Y., Narumi, T., Tanikawa, T., & Hirose, M. (2014). *Controlling fatigue while lifting objects using Pseudo-haptics in a mixed reality space. Proceedings of the HAPTICS 2014 Haptics Symposium*. New York: IEEE Computer Society. doi: 10.1109/HAPTICS.2014.6775451

Tajadura-Jiménez, A., Banakou, D., Bianchi-Berthouze, N., & Slater, M. (2017). Embodiment in a child-like talking virtual body influences object size perception, self-identification, and subsequent real speaking. *Scientific Reports*, *7*, 1. doi:10.1038/s41598-017-09497-3

Tajadura-Jiménez, A., Basia, M., Deroy, O., Fairhurst, M., Marquardt, N., & Bianchi-Berthouze, N. (2015). *As light as your footsteps. Proceedings of the CHI 2015 Conference on Human Factors in Computing Systems*. New York: ACM. doi: 10.1145/2702123.2702374

Tieri, G., Tidoni, E., Pavone, E. F., & Aglioti, S. M. (2015). Mere observation of body discontinuity affects perceived ownership and vicarious agency over a virtual hand. *Experimental Brain Research*, *233*(4), 1247–1259. doi:10.1007/s00221-015-4202-3

Tsakiris, M., Carpenter, L., James, D., & Fotopoulou, A. (2010). Hands only illusion: Multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Experimental Brain Research*, *204*(3), 343–352. doi:10.1007/s00221-009-2039-3

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: Visuotactile integration and self-attribution. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(1), 80–91. doi:10.1037/0096-1523.31.1.80

Tsakiris, M., Prabhu, G., & Haggard, P. (2006). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and Cognition*, *15*(2), 423–432. doi:10.1016/j.concog.2005.09.004

Tsakiris, M., Schütz-Bosbach, S., & Gallagher, S. (2007). On agency and body-ownership: Phenomenological and neurocognitive reflections. *Consciousness and Cognition*, *16*(3), 645–660. doi:10.1016/j.concog.2007.05.012

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: Experiencing control over the movements of others. *Journal of Personality and Social Psychology*, *86*(6), 838–848. doi:10.1037/0022-3514.86.6.838

Witt, J. K., Proffitt, D. R., & Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 880–888. doi:10.1037/0096-1523.31.5.880

Wolpert, D., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*(11), 1212–1217. doi:10.1038/81497

Yee, N., & Bailenson, J. (2007). The proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, *33*(3), 271–290. doi:10.1111/j.1468-2958.2007.00299.x

Yuan, Y., & Steed, A. (2010). *Is the rubber hand illusion induced by immersive virtual reality? Proceedings of the IEEE VR 2010 Conference on Virtual Reality and 3D User Interfaces*. New York: IEEE Computer Society. doi: 10.1109/VR.2010.5444807

Zeller, D., Friston, K. J., & Classen, J. (2016). Dynamic causal modeling of touch-evoked potentials in the rubber hand illusion. *NeuroImage*, *138*, 266–273. doi:10.1016/J.NEUROIMAGE.2016.05.065

Zopf, R., Harris, J. A., & Williams, M. A. (2011). The influence of body-ownership cues on tactile sensitivity. *Cognitive Neuroscience*, *2*(3–4), 147–154. doi:10.1080/17588928.2011.578208

Zopf, R., Polito, V., & Moore, J. (2018). Revisiting the link between body and agency: Visual movement congruency enhances intentional binding but is not body-specific. *Scientific Reports*, *8*(1), 196. doi:10.1038/s41598-017-18492-7