Multi-Destination Beaming: Apparently Being in Three Places at Once through Robotic and Virtual Embodiment

Sameer Kishore[1], Xavi Navarro Muncunill[1], Pierre Bourdin[1], Keren Or-Berkers[2], Doron Friedman[2] and Mel Slater[1,3,4]*

[1] Event Laboratory, Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain, [2] Interdisciplinary Center Herzliya, Sammy Ofer School of Communication, Herzliya, Israel, [3] Department of Computer Science, University College London, London, UK, [4] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

It has been shown that an illusion of ownership over an artificial limb or even an entire body can be induced in people through multisensory stimulation, providing evidence that the surrogate body is the person's actual body. Such body ownership illusions (BOIs) have been shown to occur with virtual bodies, mannequins, and humanoid robots. In this study, we show the possibility of eliciting a full-BOI over not one, but multiple artificial bodies concurrently. We demonstrate this by describing a system that allowed a participant to inhabit and fully control two different humanoid robots located in two distinct places and a virtual body in immersive virtual reality, using real-time full-body tracking and two-way audio communication, thereby giving them the illusion of ownership over each of them. We implemented this by allowing the participant be embodied in any one surrogate body at a given moment and letting them instantaneously switch between them. While the participant was embodied in one of the bodies, a proxy system would track the locations currently unoccupied and would control their remote representation in order to continue performing the tasks in those locations in a logical fashion. To test the efficacy of this system, an exploratory study was carried out with a fully functioning setup with three destinations and a simplified version of the proxy for use in a social interaction. The results indicate that the system was physically and psychologically comfortable and was rated highly by participants in terms of usability. Additionally, feelings of BOI and agency were reported, which were not influenced by the type of body representation. The results provide us with clues regarding BOI with humanoid robots of different dimensions, along with insight about self-localization and multilocation.

Keywords: robotics, embodiment, virtual reality, multilocation, commercial robots and applications

INTRODUCTION

Is it possible to be in more than one place at the same time? According to second century philosopher Iamblichus, Pythagoras was reported to be in two different places separated by a large distance, seen having a discussion with his disciples in both places simultaneously (Iamblichus and Taylor, 1926). Not limited only to Greek philosophy, multiple instances of bilocation (and even multilocation) have

been reported over the centuries as part of folklore, mysticism, and in various religious texts, with most examples considered to be miraculous or divine events. However, none of these alleged events have any scientific basis, since obviously one person being in multiple places simultaneously is physically impossible. Today, although through the transformation of aspects of reality into the digital domain, it is possible to apparently violate the laws of physics – to be evidently present in two (or more) different places at the same time and be able to interact with people and carry out tasks at all of them. Already, attempts have been made to virtually manipulate certain constraints of physical reality by simulating time travel in immersive virtual reality (Friedman et al., 2014). Similarly, although being in multiple locations simultaneously is physically impossible, it can be simulated partially by leveraging the concept of telepresence.

The concept of bilocation, in terms of presence with respect to virtual environments, has been studied in the past (Gooskens, 2010), and it has been shown that humans can distribute their "self" over two distinct places (Wissmath et al., 2011). In their study, participants experienced a virtual rollercoaster and were continuously asked about the extent to which they felt present in either their real location or the displayed virtual location. Their results showed that the participants were successfully able to distribute their self over these two distinct places, with the sum of their "divided" self adding up to approximately 100% at all times. However, we have found no studies that have attempted to study the phenomenon over three or four places. Several telepresence applications have been developed that allow operation *via* remotely controlled robots, for example, to perform "telesurgery" (Marescaux et al., 2001) or a semi-autonomous teleoperated humanoid robot in space (Bluethmann et al., 2003). Furthermore, there exist numerous "telepresence" robots that instead of solving a highly specific problem provide a more general-purpose utility such as for social interactions (Guizzo, 2010; Tsui et al., 2011). These robots allow users to move around in a remote space and interact with people present there. However, these systems do not usually have a high quality of interaction, as most traditional teleoperation and telepresence systems are desktop-based, with control of the remote robot limited to using joysticks or other traditional forms of input. The view of the remote location in a majority of the systems is presented through a standard 2D display. However, there have been studies where participants have control over the remote robotic representation using natural movements and gestures (Adalgeirsson and Breazeal, 2010). Furthermore, the concept of telepresence has been studied with humanoid robots as well where two distinctly located humanoid robots were used as surrogate representations to explore the idea of bi-directional telepresence (Nagendran et al., 2015). Although the participants were not able to see the other person or the remote location, they were able to communicate verbally and through gestures.

Although these systems are effective in eliciting the feeling of telepresence, they allow operators to control only one remotely located physical entity at a time, and thereby experience presence in only one remote location. Although there has been a system that lets an operator monitor up to four robots in different physical locations (Glas et al., 2008), one key feature is still lacking, regardless of the number of remote teleoperators. This is the crucial feeling elicited in the operator of the illusion that the robot body is their body (body ownership) rather than only remote presence through control of and sensing *via* the robot. Also, the local people with whom the interaction takes place should ideally have a strong sensation of actually being with the embodied person. Thus, by exploiting the concepts of telepresence, teleoperation, and embodiment in a robot body and having multiple remote locations, each equipped with a remote surrogate body that the participant can feel ownership of, the phenomenon of bilocation or multilocation can be simulated.

The concept of body ownership has been studied extensively, and most relevant to our discussion is a case where a mannequin body was used to substitute the real body of participants (Petkova and Ehrsson, 2008). This was implemented by using stereo cameras mounted on the mannequin's head and a head-mounted display (HMD) streaming the video from the first-person perspective of the mannequin. By giving participants visual feedback that their real body had been replaced by a mannequin body and providing synchronous tactile simulation, the experiment was able to elicit the illusion of full-body ownership over the mannequin body. This phenomenon has been successfully replicated several times in virtual reality (Slater and Sanchez-Vives, 2014) and various aspects of the illusion have been tested, such as varying the racial appearance of the virtual body (Kilteni et al., 2013; Maister et al., 2013; Peck et al., 2013), limb size (Kilteni et al., 2012), body size (Normand et al., 2011), and also by modifying the age of the virtual body (Banakou et al., 2013). Over several studies, certain factors have been found to have a major effect on the illusion, such as first-person perspective, and the body having a human morphology (Maselli and Slater, 2013). Additionally, giving participants a natural way of controlling the body can also strongly increase feelings of agency over the body (Tsakiris et al., 2006). One way of achieving agency is to have full-body tracking of all movements produced by the participant and mapping these to the corresponding limbs of an avatar or a robot in real time. This has been applied to generation of a body ownership illusion (BOI) over a humanoid robot body seen from first person perspective (Kishore et al., 2014), using the algorithm described in Spanlang et al. (2013). Methods have also been developed to modify the appearance of a humanoid robot, by projecting a 3D model of the head of the participant on the face of the remote robot using the shader lamp method (Raskar et al., 2001). The visitor could be far away seeing the real surroundings of the physical surrogate (in this case an animatron) through a virtual reality, and his or her face back-projected onto a shell, so that an observer of the surrogate would see video of the real face of the distant person and be able to interact with him or her (Lincoln et al., 2011).

Thus, combining these concepts of telepresence, teleoperation, and embodiment can give rise to simulated "teleportation" where participants have the illusion of being in a remote location, with a physical body that they feel as their own, which they can use to sense the environment, to explore the remote location, and to interact with the people there. Several studies have attempted to investigate this phenomenon and develop applications where such a system could be applied, such as performing remote acting rehearsals, instruction of a musical instrument, remote medical rehabilitation (Normand et al., 2012b;

Perez-Marcos et al., 2012), and even interspecies communication where a human was embodied as a small robot interacting with a rat in a cage (Normand et al., 2012a). In another recent application developed using a similar concept, a journalist was able to embody a humanoid robot located in a remote destination and carry out interviews with people there (Kishore et al., 2016). All the applications mentioned above follow a similar setup in terms of the various components that form part of what has been called a "beaming" system, which consists of the following system components (Steed et al., 2012):

- *Visitor*: the person who "travels" to the remote locations is referred to as the visitor.
- *Locals*: the locals are the people present in the various remote destinations. The locals can interact with each other and the visitor – *via* the visitor's remote representation.
- *Transporter*: the transporter is the physical location of the visitor who uses the technology to "travel" to remote places. The transporter would be equipped with display and tracking devices. For example, for display a HMD, headphones and haptic devices would be used. For tracking, there would be full-body and head tracking systems and a high quality microphone for voice input. The display devices are for immersion of the visitor in the remote environment, whereas the tracking systems are for controlling the remote representation.
- *Destination*: the destination is the remote location to where the visitor virtually travels. The destination can be a physical location or even a shared virtual environment. The important factor in either case is for the visitor to be represented in some way at the destination.

Although the applications mentioned above provide a highly immersive system for remote interaction, they are all designed for a single remote destination. Here, we describe a system that builds upon this idea by introducing the novel concept of multiple concurrent remote destinations, each visited using a remote (robotic or virtual) body that can be inhabited and controlled by the visitor, thereby giving them the illusion of ownership over each of them. Due to its modular nature, the system implements this not only by immersing and embodying participants in one destination at any moment in time but also allowing them to instantaneously switch between the various destinations.

While a participant is present in one destination, it is crucial to have a method for engaging the locals in the remote destinations where currently the visitor is not embodied. Hence, a proxy system has also been developed, which tracks the destinations not currently inhabited by the participants, and takes over their remote representation in order to continue performing the tasks in those locations, so that the locals ideally do not realize that the visitor is not currently present. The proxy was developed based on the system described in Friedman et al. (2013) and Friedman and Hasler (2016).

To test the efficacy of this system, an exploratory study was carried out with a fully functioning setup with three destinations and a simplified version of the proxy for use in a social interaction. Participants could be in three distinct remote locations (two physical and one virtual) concurrently, where they were in full control of a different humanoid robot at each physical location and an avatar in the virtual location. They could see, hear, and interact with the people present in all three remote locations and had to perform predefined tasks.

MATERIALS AND METHODS

The system has three main components – the transporter, the destinations, and the proxy.

The Transporter

The transporter system is at the physical location of the person to be "beamed." Thus, the most important task of the transporter is to capture the multisensory data of the visitor and stream it to the destination in real time. This includes, but is not limited to capturing all the translation and orientation data of each limb of the visitor, and the audio stream. The complementary objective involves recreating the environment of the destination for the visitor, as accurately as possible. This requires displaying the view of the destination from the perspective of the visitor's representation, including an audio stream to enable interaction with the locals. Furthermore, the transporter also needs to keep track of the destination that the visitor is currently in and to switch to the other destinations when required. The proxy system, which manages the unoccupied destinations in an appropriate manner, is also associated with the transporter and is discussed in detail below.

In order to implement a system that captures and streams the visitor's behavior to the surrogate at the destination, several aspects need to be considered. To capture all position and orientation data of the visitor's limbs, a full-body tracking system is used. Several full-body tracking systems such as the Xsens Inertial Motion Capture System[1] and the Microsoft Kinect[2] were successfully incorporated and tested. For the experiment described below, we used the Arena Optitrack[3] system, which works with a suit that contains markers tracked by 12 infrared cameras fitted in a room. Although any of these tracking systems could be used, choosing the appropriate one depends on the type of setup. Since we carried out a study where multiple participants would be using the system in a single day, flexibility of the system to adapt to different people was of the utmost importance. For this reason, the Xsens system was not used as it typically takes up to 2 h to rewire a suit of a different size. On the other hand, this can be easily done in the Optitrack system as the process involves simply replacing the reflective markers on to another suit *via* velcro. In terms of portability of the transporter setup, the Microsoft Kinect would be a better choice than the Optitrack system due to the ease of setting up the Kinect; however, in our case, portability was not a concern since the study was carried out inside the laboratory.

The head of the visitor is tracked separately using a 6-dof Intersense 900.[4] This is done since the Intersense 900 is highly accurate and has a low latency of 4 ms, which allows high quality capture of the participant's head movement. All these

---

[1]https://www.xsens.com/products/xsens-mvn/.
[2]https://www.microsoft.com/en-us/kinectforwindows/.
[3]http://www.optitrack.com/products/motive/body/.
[4]http://www.intersense.com/pages/20/14.

captured data are then converted and mapped in real time to either a humanoid robot or a virtual avatar depending on the type of destination in which the visitor is currently present. This is achieved using a library that was developed specifically for providing natural method of teleoperation of avatars and robots (Spanlang et al., 2013). Since the visitor has the ability to move the body of the robot or avatar by simply moving their own body, this is likely to result in a high sense of agency over the remote body. This phenomenon is a strong factor in inducing the full-BOI (Kilteni et al., 2015).

Complementary to the capture and streaming of the visitor's behavior, the environment of the destination also has to be recreated at the transporter. In order to provide the visitor with a view of the destination, a fully tracked wide field-of-view HMD was used. The HMD used for the experiment was an NVIS nVisor SX111,[5] on to which was attached the Intersense 900 tracker mentioned above, and the system has also been tested with the Oculus Rift.[6] The NVIS was chosen for this study due to its high resolution (1280 × 1024 per eye) and high field-of-view [76°H × 64°V field of view (FOV) per eye in a 13° outward rotated setup, total horizontal FOV 111°] as compared to the Oculus Rift DK1. However, a major concern while using this HMD is its weight of 1.3 kg, which may cause discomfort if worn for a prolonged period of time. In the case of this study the total time of the experience was under 10 min, but if this had been longer, the Oculus Rift would have been the better choice. If the visitor is present in the virtual destination, they are provided with the view of an immersive virtual environment with avatar representations of themselves and the locals. They have full control over the limbs of the avatar, and they can look down to see the avatar body that they currently occupy collocated with their real body, which cannot be seen through the HMD. They can move around in the virtual environment and interact with the locals *via* physical gestures and a two-way audio stream. In case of the physical destinations, the visitor views the environment through the "eyes" of the humanoid robot in which they are embodied. In our experimental study, the two robots present at the two destinations were each fitted with a set of two cameras on the head, separated by the standard interocular distance that provided a stereoscopic 3D video stream *via* the Internet directly to the HMD of the visitor. Since the cameras were placed on the head of the robot, the view of the destination as seen through the cameras can be updated by the visitor by simply moving their own head.

In order to provide a high quality, immersive experience to the visitor and the local, it is essential to synchronize the various data streams (video, movement, and audio) and minimize the observable latency as much as possible. Thus, head tracking and full-body tracking systems were chosen for their high accuracy and low latency in terms of capturing and streaming the data. All the streams of data were transmitted through a high-bandwidth network connection. In terms of bandwidth usage, the stereoscopic video streams required the most amount of bandwidth.

Thus, the video frames were compressed using the VP8 video codec prior to streaming (Bankoski et al., 2011). Finally, given the complexity of our setup, it is quite difficult to measure latency of all the data streams precisely; however, it was observed during pilot studies to be low enough for the system to work successfully.

The Destinations

In the case of this study, there are two physical destinations where the visitor was represented by a different humanoid robot and by an avatar in the virtual destination. Since the aim of the system is to allow the visitor and locals to have a natural interaction, the destinations and the locals themselves should be fitted with the least intrusive equipment possible. Two types of possible destinations have been implemented in the system – physical and virtual. A physical destination is an actual remote location, while a virtual destination is a collaborative immersive virtual environment that the visitor and locals share. For the sake of consistency, we will refer to the two physical destinations as PD1 and PD2, while the virtual destination will be referred to as VD3.

Physical Destinations

A major factor that dictates the setting up of a destination is that the interaction between the visitor and locals has to be natural, and not constrained by intrusive equipment and cables. Thus, the only essential component in the physical destinations is the humanoid robot that represents the visitor. The robot used for PD1 was the Robothespian, manufactured by Engineered Arts, UK. The robot is 180 cm tall biped humanoid robot with pneumatic arms, and a DC motor for the head and torso movement. The DC motors provide 3 degrees of freedom each for the head and torso. For streaming the view of the destination, the robot has two Microsoft HD-3000 webcams on its forehead. On the other hand, PD2 was equipped with a Nao robot, manufactured by Aldebaran Robotics, France. The Nao is also a humanoid robot, although is much smaller in height at approximately 60 cm. All the joints are operated by DC motors, with 2 degrees of freedom in the head and 5 degrees of freedom in each arm and leg. A helmet with two Logitech webcams separated by the standard interocular distance is mounted on top of the robot's head.

The cameras mounted on the heads of the robots can lead to a stronger feeling of agency and head-based sensorimotor contingencies as the view of the cameras changes based on the visitor's head movement. Furthermore, looking down and seeing the surrogate robot body instead of their own increases the feeling of the BOI. However, participants could not look down toward their body and see it substituted by the robot body due to the limitations of the range of the head movements of both robots. To overcome this, a mirror was placed at both PD1 and PD2 pointed toward the robots, as this has been shown to induce stronger feelings of the BOI (González-Franco et al., 2010; Preston et al., 2015).

Virtual Destination

One important property of a virtual destination is that it can be completely customized based on requirements of the task to be carried out. It could be indoors or outdoors and with as many avatars present as the number of locals in the scenario. VD3

was designed as a room with one avatar experienced from first person perspective representing the visitor, and another seated avatar as the local. The visitor experienced visuomotor synchrony with respect to the movements of their avatar. The entire virtual environment was programed in the Unity3D software.

Two-way audio was also implemented between the transporter and whichever destination the visitor was in at any given point in time. An underlying server was developed that would be able to detect and manage the currently occupied destination of the visitor. The server would keep track of the destination the visitor was currently in and would connect their audio streams together, while letting the proxy audio take over at the other destinations. Similarly, the movements of the visitor would also be streamed only to their current destination, with the proxy system controlling movements of the other representations.
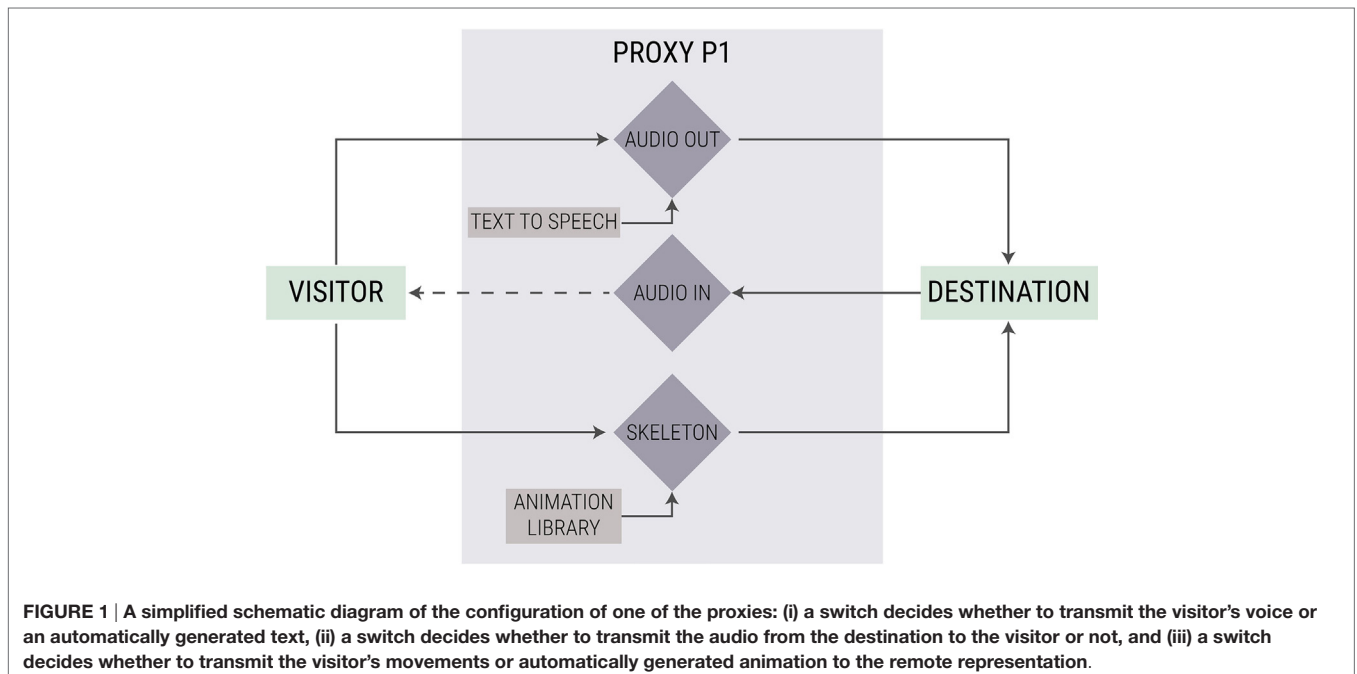
The Proxy System

A critical feature that gives the visitor the ability to be apparently in three different places simultaneously and perform tasks at each destination is the proxy system. While the visitor is at one destination, the other two destinations are running concurrently as well, where an artificial agent referred to as the proxy continues interacting with locals and attempts to perform tasks, on behalf of its owner. The visitor has the option to switch to any of the other destinations instantaneously and take over the robot or avatar representation in that destination. At the instant of the switch, the proxy hands control over to the visitor and takes over the representation that the visitor has just left in the previous destination. In order to enable the visitor to be in those three places at the same time, we launch three instantiations of the proxy: P1, P2, and P3, correspondingly.
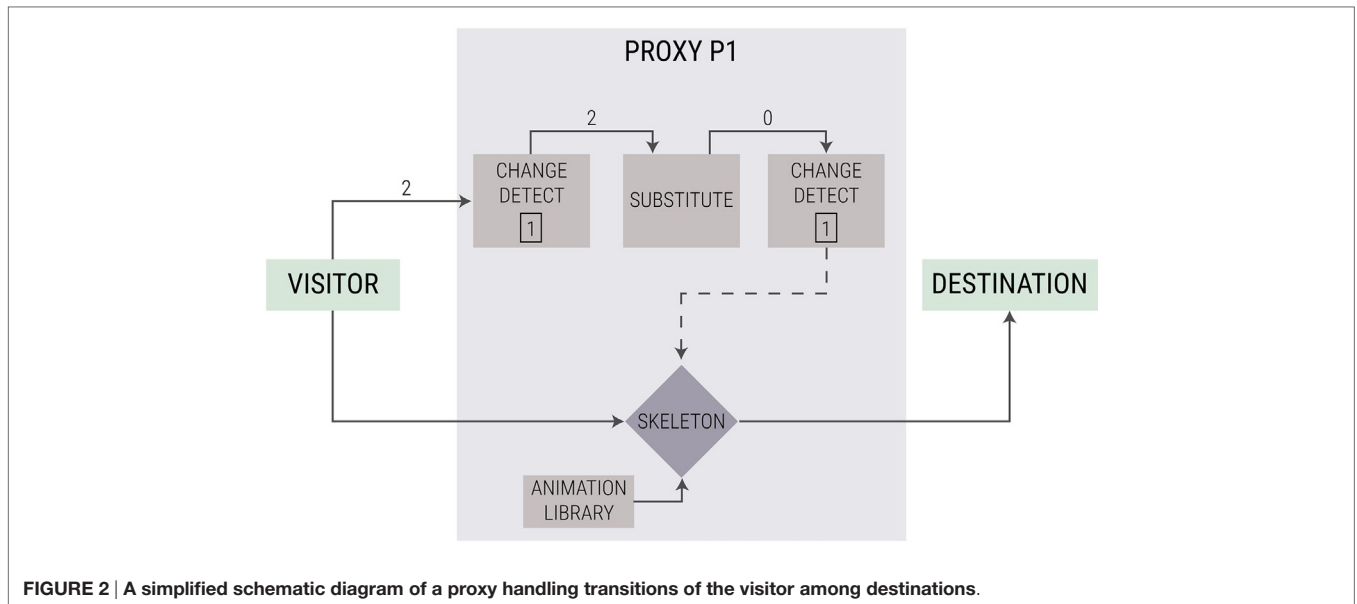
Each proxy is configured to allow a transition of control from the visitor to the proxy and back, as in **Figure 1**. The proxy has

two audio channels: one for receiving and one for sending. These channels are always open. In addition, the proxy has one channel for skeleton animation that controls the non-verbal behavior of the representation. Other than these two data streams, the proxies' functionalities are dependent on processing application-dependent events, for each one of the three destinations. The three proxies are always "alive" and generate output streams. At each point in time, one of them is ignored and is overridden by the visitor. In principle, we could have also opted to pause some of the computations made by that proxy when the visitor takes control and restart them when it needs to take control back, but in the case of this straightforward setup, there would have been no computational advantage in doing this.

The transition of the visitor between the destinations is handled by a generic mechanism that is implemented in all proxies identically. In general, the proxy's architecture is based on generic components (Friedman and Hasler, 2016); this is intended to easily obtain a wide range of proxy behaviors by using only configuration, rather than programing. The switching implementation is achieved by configuring the "change detection" and "substitute" components, as in **Figure 2**. In this figure, we consider, for example, the case of proxy 1. Let us assume that the application is initiated such that the visitor is in destination 1, and thus the proxy is in background mode (the owner is "in the front" and controls the representation). The fact that the visitor is in destination 1 is maintained as a "long-term memory" by the two "change detection" components.

In the figure, we illustrate the moment at which the visitor decides to switch to destination 2. The first change detection component detects that there is a change ($1 \neq 2$) and sends the information onward. Next, the information is handled by the "substitute" component. This component is used in our case as follows: if the visitor is now in the destination that the proxy

FIGURE 1 | A simplified schematic diagram of the configuration of one of the proxies: (i) a switch decides whether to transmit the visitor's voice or an automatically generated text, (ii) a switch decides whether to transmit the audio from the destination to the visitor or not, and (iii) a switch decides whether to transmit the visitor's movements or automatically generated animation to the remote representation.

is responsible for, then pass on that number, otherwise replace it by "0." In general, this component works as a "search and replace" action: it is able to detect patterns of text (using regular expressions) and replace them under certain conditions. In this case, the input to this component is "2." Since this is proxy 1, the component replaces the input "2" by an output "0" (indicating "the visitor is not in our destination").

The input from the "substitute" component is fed into the second "change detection" component. In our example, this would also trigger a change: earlier the visitor was here (1) but now it receives a "0." Thus, it triggers an event to the switches. The switches work as on/off toggles. As before they were set to pass on information from the visitor, they now switch to passing on the information automatically generated by the proxy. Using this configuration, based on reusable generic components, all proxies have the same mechanism, and thus the same proxy configuration can be used arbitrarily in different numbers of destinations.
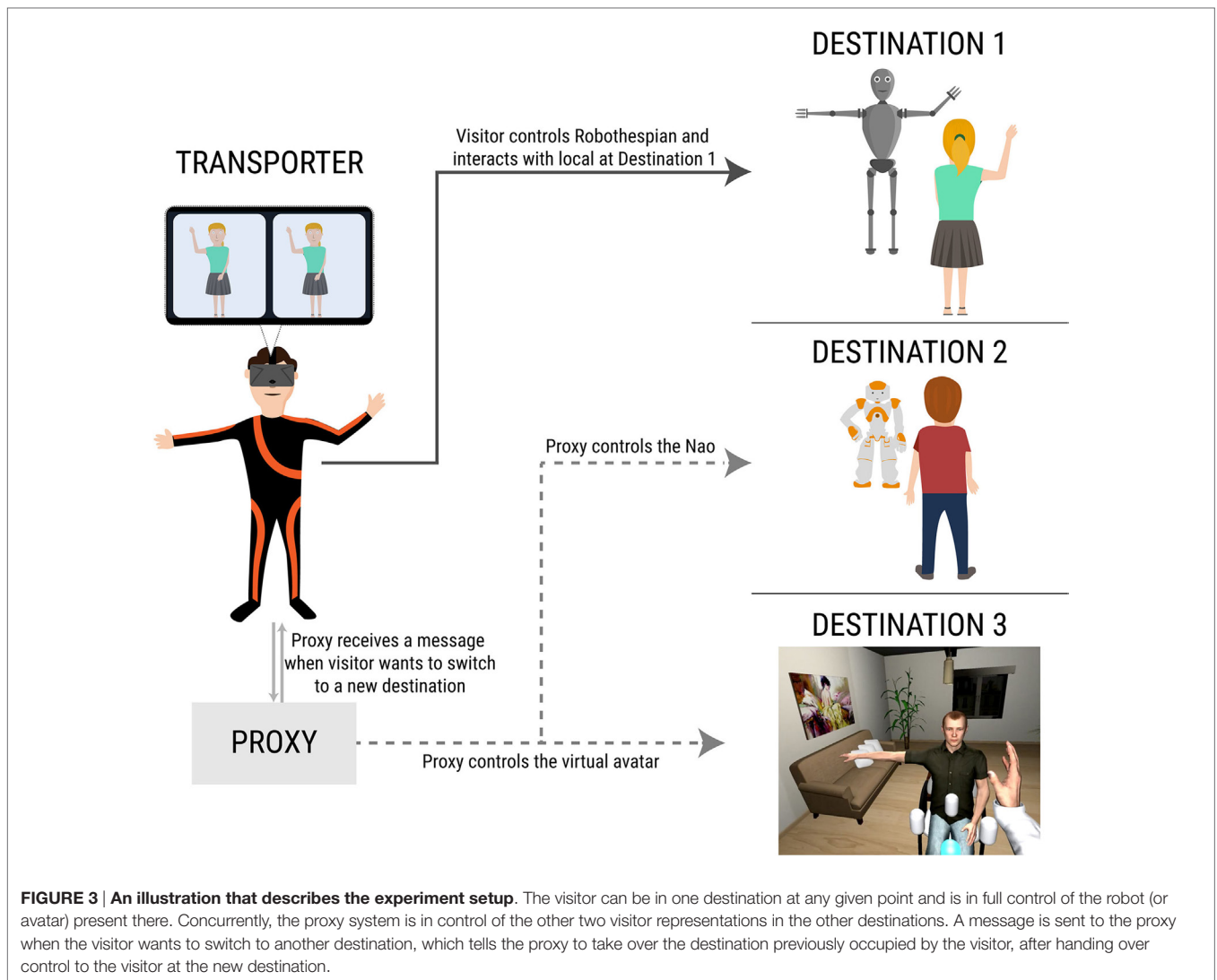
Experimental Setup

In order to test the efficacy of the system, an experiment was designed with two main objectives in mind. The primary question addressed by the experiment was whether the system could be used successfully and efficiently as a tool for interaction by a participant apparently present at three remote locations simultaneously. The subsequent question concerned the quality of the experience. Specifically, we explored the possibility of eliciting a full-BOI in the remote representations, and whether such an illusion assists in creating a stronger experience of being physically present in the same location as the locals. Furthermore, to test the merit of the proxy system, it was important to observe the interaction from the perspective of the local as well.

A total of 41 participants (mean age = 27, SD = 10.5, 17 males, 24 females) were recruited, with each session involving 2 participants – 1 "visitor-participant" and 1 "local-participant" (21 visitors; 20 locals); these will be referred to simply as the visitor or

local for the remainder of the paper. The last session was run with a confederate instead of a local participant. The participants were recruited around the campus by advertisement and paid for their participation. This study was approved by the Comissió Bioètica of the University of Barcelona.

Each session involved the visitor "beaming" to three destinations, two physical and one virtual. As mentioned earlier, PD1 was equipped with the Robothespian, PD2 with the Nao, and VD3 with an avatar. The experiment was designed for social interactions, and the tasks were designed accordingly. The visitor would start with PD2 in all cases and would have a casual, friendly conversation with the local there. For this reason, each visitor and local were recruited in pairs, so they knew each other beforehand, and would be able to talk to each other comfortably. Next, at PD1, the visitor would have to give a talk on any topic that they preferred. They were told beforehand to prepare a 5-min talk on any subject they liked. At PD1, a confederate was present to listen to the talk. In VD3, the visitor would have to help a virtual avatar that was seated in front of them to perform a simple exercise with their right arm. The local-avatar at VD3 was controlled by the experimenter, as the task was straightforward and there was no scope for further interaction other than performing the exercise. The main reason for adding the virtual destination was to observe whether the visitor would be able to cope with "being present" at three destinations – with one of them being virtual. **Figure 3** shows an illustration of the entire setup, also shown in Video S1 in Supplementary Material (and at https://youtu.be/oh1B6C3JggQ).

Each session followed the same sequence – the visitors would begin at PD2 with a friendly conversation with the local. After 2 min, they would automatically be switched to PD1, and then 2 min later to VD3. After that, the cycle would repeat again in the same order and the visitors would visit each destination the second time. While in principle our system allows the visitors to themselves choose when to switch between destinations, for experimental control we opted for automatic switching after a

FIGURE 3 | An illustration that describes the experiment setup. The visitor can be in one destination at any given point and is in full control of the robot (or avatar) present there. Concurrently, the proxy system is in control of the other two visitor representations in the other destinations. A message is sent to the proxy when the visitor wants to switch to another destination, which tells the proxy to take over the destination previously occupied by the visitor, after handing over control to the visitor at the new destination.

fixed amount of time, since all tasks were open-ended and thus needed a time limit. While the visitor was at PD1 and VD3, the local participant at PD2 would continue interacting with the Nao humanoid robot, at that point being controlled by the proxy. The main reason behind visiting each destination the second time was to show the visitor that the interaction had continued even after they had left. **Figure 4** shows the Transporter and each of the three destinations. For the purpose of the experiment, a simplified version of the proxy was used, the working of which is explained below.

Before starting the experiment, the visitor was asked to record a few generic sentences such as:

- "*Oh! I forgot what I was saying!*"
- "*I really can't remember what I was going to say. This has never happened before, you know!*"
- "*By the way, my favorite movie is <Visitor's favorite movie>. What's yours?*"
- "*Anyway, forget about me, what's up with you?*"

If the proxy would detect that the visitor was not in a certain destination, it would play back the statements in random order and at random times. Additionally, pre-recorded "idle" animations were also loaded in the Animation Library of the proxy system, which would be streamed to the unoccupied robot or avatar. Before beginning the experiment, the visitor was given a brief description of the system and was informed about the tasks that they would have to perform. They were also told that they would be automatically switched from one destination to the next after 2 min. The local was not informed of this otherwise they would expect the proxy to take over, which would break the interaction. Both the participants (visitor and local roles) were interviewed and given questionnaires after the experience. An ongoing Multi-Destination Beaming session can be seen in **Figure 5**.

Measures

Our main focus was to investigate whether the system was effective in allowing the visitor to perform tasks in three distinct places simultaneously. The question of usability was approached from
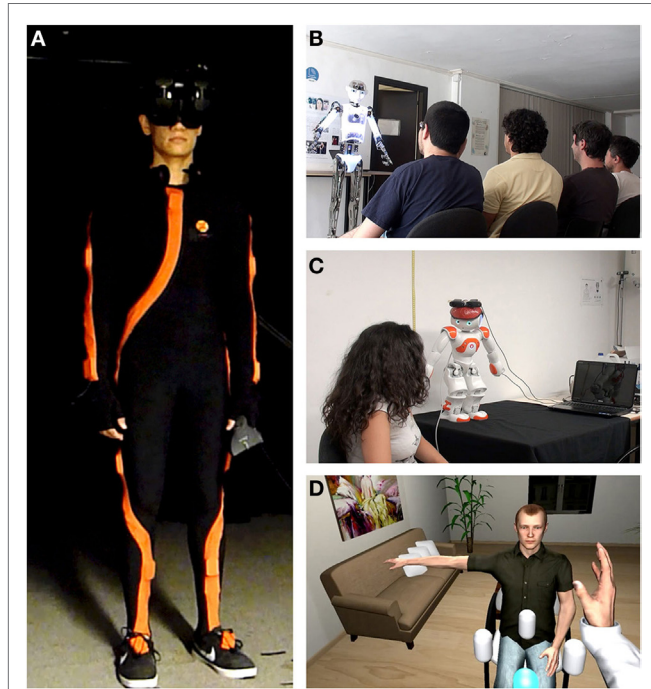
several perspectives, which are detailed as follows. Visitors were asked to assess their feeling of presence, specifically place illusion (PI) (Slater, 2009) and degree of success with which they could cope with the tasks as well as their physical and psychological comfort level while using in the system, along with their overall rating of the system's usability (**Table 1**). In addition to the questions related to task performance, the visitors were also asked

FIGURE 4 | **Images from the transporter and all three destinations.** Clockwise from left: **(A)** the visitor wears the Xsens Inertial motion capture suit and the NVIS nVisor SX111 HMD at the transporter; **(B)** PD1 where the visitor gives a talk to four locals, while embodied in the Robothespian; **(C)** PD2 where the visitor is in control of a Nao and engages in a friendly, social conversation with a friend; **(D)** VD3 where the visitor assists a local in performing an exercise, while embodied in a virtual avatar.

questions related to the BOI and agency twice, once for each robot (**Table 2**). The embodiment questionnaire, designed based on a previous study that was carried out using the same robot (Kishore et al., 2014), was not given for VD3 since high ownership illusions and high levels of agency have been reported in various previous studies using the exact same setup (Banakou et al., 2013; Kilteni et al., 2013; Peck et al., 2013). The locals were asked two questions regarding their experience (**Table 3**). Questionnaires were administered in paper form immediately after the experiment, and in the same laboratory where the experiment took place. All responses were on a seven point Likert scale anchored with "Strongly disagree" (1) and "Strongly agree" (7). After finishing the questionnaire, the experimenters conducted an oral interview with both participants as well.
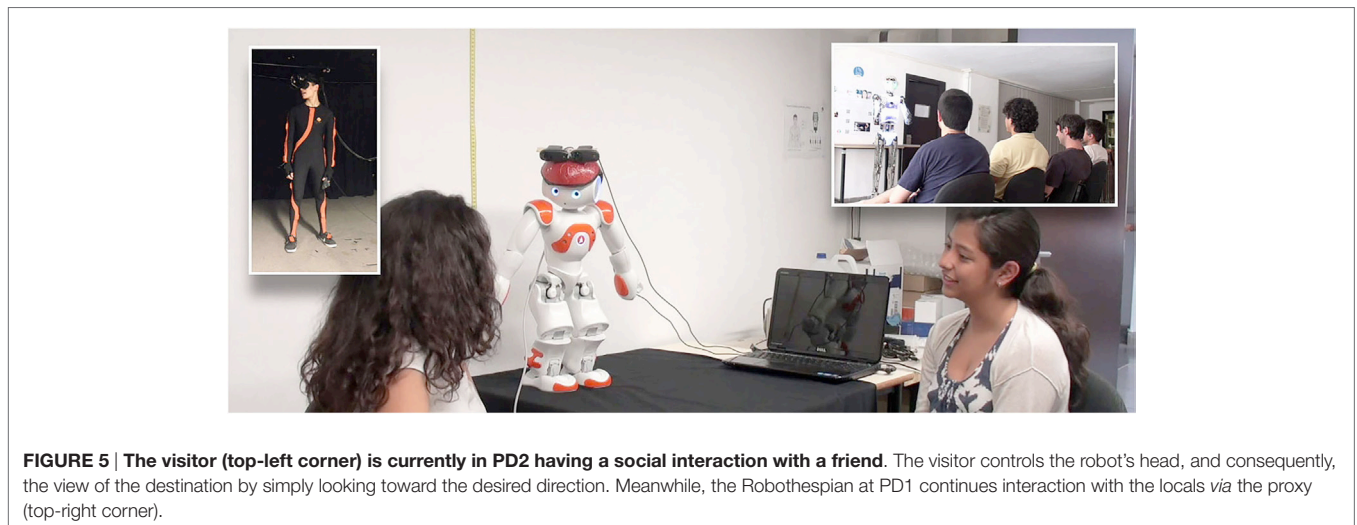
The question of full-BOI over the two robots was addressed through questions *MyBody*, *MirrorBody*, *Arm_Mine*, *TwoBodies*, and *Machine*, agency was assessed *via Agency* and *ControlBody* while PI was assessed by *PlaceIllusion* and *Copresence*. The question *ArmMine* was asked since the arms were the only part of the robot that was visible directly by looking down. Task performance ratings were measured using questions *CopeTasks* and *AccomplishTasks*. Finally, questions related to overall usability and comfort level of the system were *PsychologicalComfort*, *PhysicalComfort*, and *Usability*. The experience of the local was evaluated through two specific questions: similar to the visitor, *Copresence* was used to assess the local's feeling of being with another person in the same space. *Interact_Local* was used in order to evaluate performance of the proxy system.

RESULTS

The responses of the visitors to the questions in **Tables 1** and **2** are shown in **Figures 6** and **7**. Based on the responses to the question *CopeTasks*, the system was effective in giving participants the feeling that they could cope with the tasks (*median* = 6, *IQR* = 2) regardless of the destination, and that they could also accomplish the tasks (*AccomplishTasks*) (*median* = 6, *IQR* = 1.75). The high scores for the question *CopeTasks* for all

FIGURE 5 | **The visitor (top-left corner) is currently in PD2 having a social interaction with a friend.** The visitor controls the robot's head, and consequently, the view of the destination by simply looking toward the desired direction. Meanwhile, the Robothespian at PD1 continues interaction with the locals *via* the proxy (top-right corner).

TABLE 1 | Questions related to PI, copresence, task performance, comfort, and system usability.

Copresence	How much did you feel you were physically with the other people?
PlaceIllusion	How much did you feel you were in the places that were displayed?
CopeTasks	How much did you feel you were able to cope with the tasks?
AccomplishTasks	How much did you feel you were able to accomplish the tasks?
PhysicalComfort	Could you rate how physically comfortable you felt while inside the system?
PsychologicalComfort	Could you rate how psychologically comfortable you felt while inside the system?
Usability	How much would you use this system if you really had to be at 3 different places at the same time?

Responses were given on a Likert scale anchored from 1 (disagree strongly) to 7 (agree strongly).

TABLE 2 | Questions related to body ownership illusion and agency.

MyBody	I felt as if the body I saw in the mirror might be my body
Agency	I felt like I controlled the robot body as if it was my own body
MirrorBody	When I looked in the mirror, I felt as if the body I saw might be my own
Arm_Mine	When I saw the robot arm moving, I felt as if the arm I saw might be my own
TwoBodies	I felt that I had two bodies
Machine	I felt I was turning into a machine
ControlBody	I felt in control of the robot body

Responses were given on a Likert scale anchored from 1 (disagree strongly) to 7 (agree strongly).
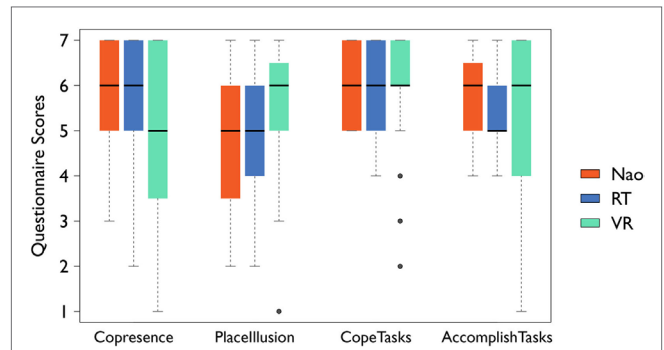
TABLE 3 | Questions related to the local participant experience.

|	
Copresence	How much did you feel you were physically with the other people?
Interact_Local	How much did you feel you were interacting with your friend?

Responses were given on a Likert scale anchored from 1 (disagree strongly) to 7 (agree strongly).

FIGURE 6 | Responses to questions on place illusion and task performance per destination, as provided in Table 2. Responses were given on a 7-point Likert scale, where 7 is "strongly agree" and 1 is "strongly disagree." Error bars represent the SE and dots are statistical outliers.
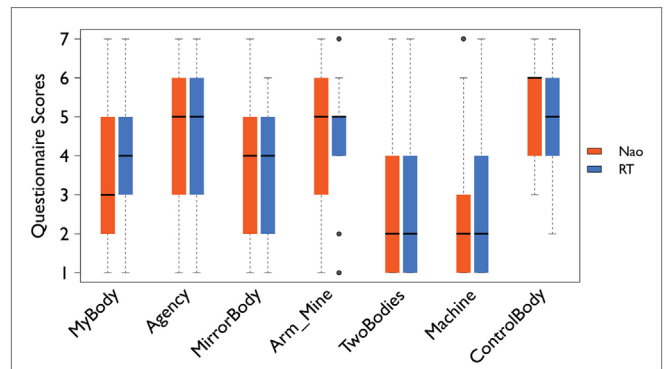
FIGURE 7 | Responses to questions on embodiment and agency per robot, as provided in Table 1. Responses were given on a 7-point Likert scale, where 7 is "strongly agree" and 1 is "strongly disagree." Error bars represent the SE and dots are statistical outliers.

three destinations suggests that the type of representation in the various destinations did not make any difference to their task performance for this particular sample of people and social interaction tasks. Moreover, comparing across destinations would not have been appropriate had the tasks had been quite different at each destination.

Participants gave quite high scores for both PI (*PlaceIllusion*) for each of the three destinations as shown in **Figure 6** (*median* = 5, 5, 6; *IQR* = 2, 2.25, and 1.25 for *PD1*, *PD2*, and *VD3*, respectively), as well as for the feeling of being in the same space as the locals (*Copresence*) (*median* = 6, 6, 5; *IQR* = 2, 2, 3.25). No significant differences were found between the responses to the copresence ratings between the physical and virtual destinations (*Wilcoxon signed-rank test: Copresence(VE) – Copresence(Nao): Z = −1.43, p > 0.05; Copresence(VE) – Copresence(RT): Z = −1.75, p > 0.05*). Similarly, the differences in responses for PI with respect to the physical and virtual destinations were not statistically significant either (*Wilcoxon signed-rank test: PlaceIllusion(VE) – PlaceIllusion(Nao): Z = −1.62, p > 0.05; PlaceIllusion(VE) – PlaceIllusion(RT): Z = −1.17, p > 0.05*). Furthermore, visitors

gave very high scores for the physical and psychological comfort felt while using the system (*PhysicalComfort: median = 6, IQR = 1.25; PsychologicalComfort: median = 6, IQR = 2*). Finally, the question regarding their overall evaluation of the usefulness of the system was also met with strongly positive responses as participants gave high scores for using the system in the real world to be in three different places at the same time (*Usability: median = 6, IQR = 1*). Similarly, the local participants also had a positive sensation of being with the visitors as well (*Copresence: median = 5, IQR = 2*) and gave positive responses to the question regarding the robot they were interacting with being a surrogate representation of their friend (*Interact_Local: median = 5, IQR = 1*).

Regarding the BOI over the two robots and sense of agency, participants gave positive responses for the relevant questions in the questionnaire and the post-experience interview as well. The differences in scores for the questions *MyBody* and *TwoBodies*, with respect to the two robots were not statistically significant (*Wilcoxon signed-rank test: MyBody(RT) – MyBody(Nao): Z = −0.159, p > 0.05; TwoBodies(RT) – TwoBodies(Nao): Z = −0.552, p > 0.05*). Moreover, *Agency* had exactly the same

scores for both robots (*median* = 5, *IQR* = 3), while *ControlBody* had a higher median score for the Nao (*median* = 6, *IQR* = 2) than the Robothespian (*median* = 5, *IQR* = 2).

DISCUSSION

The discussion of the results is presented in two parts – the first section relates to the characteristics and efficacy of the system, such as task performance, physical and psychological comfort, and the overall usability ratings of the system. Additionally, the responses by locals can also be taken into account here, since they provide insight about the ability of the proxy system and the influence of a physical representation of the visitor in the destination. The second aspect of the discussion is related to the visitor's BOI and feeling of agency over the two remote humanoid robots, and the significance of those in terms of the system's performance.

System Performance

Prior to the experiment, the main concern regarding the system had been that the visitors might find it difficult to keep up with being at three different places simultaneously, and coping with three different tasks. However, most did not express any complaint regarding this in the interviews or in the questionnaires. On the contrary, they seemed comfortable with the idea as shown by questionnaire results (**Figure 6**, *CopeTasks* and *AccomplishTasks*). It is understandable that there are no significant differences according to destination type for the variable *CopeTasks*, since the tasks were mainly related to verbally interacting with the locals. If the tasks had been more physically oriented, the difference in dimensions of the robots or lack of haptic feedback in case of the virtual destination would have been more important.

Although responses were not statistically significant, median scores of the questionnaire data related to copresence give some indication that the two physical destinations performed slightly better than the virtual destination. This is understandable since the visitors interacted with real locals in PD1 and PD2, while the virtual local was automated in VD3, instead of being a representation of a real person. On the other hand, scores for PI are higher for the virtual destination as compared to the two physical destinations, which could be attributed to the higher resolution and higher frame rate of the rendered environment as compared to the lower quality video stream. It has been shown that display parameters such as frame rate have an effect on the degree of PI (Barfield and Hendrix, 1995; Sanchez-Vives and Slater, 2005). Additionally, visitors had much greater freedom of movement in the virtual environment embodied as the avatar, while they were constrained by the limited range of movement of the robots in the physical destinations. The responses in the post-experience interviews were mostly positive, with comments such as: "*I was in 3 different places with 3 different people*" and "*I felt transported to 3 different destinations, I had conversations with 2 individuals and prompted physical exercise to a patient.*" Furthermore, regarding the moment they were transported from one destination to the next, most participants answered that they could cope with this very easily, even though for this experiment the switching was done automatically. We anticipate that in actual applications this would typically be under the visitor's control. The transition from

one destination to another was instantaneous, without any fading or blurring, which allowed visitors to immediately resume control over the surrogate representations. A participant commented, "*I felt like switching from one environment to another was seamless – it didn't feel weird or artificial.*"

In addition, it can also be said that the system was comfortable enough to use for the purpose of performing tasks in three locations, given the high scores to the questions related to comfort and usability. These high scores for physical comfort are in spite of the fact that the HMD used for the system weighed 1.3 kg and the participants had to don a full-body tracking suit before the start of the experiment.

Furthermore, local participants also gave high scores to the question related to copresence (*Copresence: median* = 5, *IQR* = 2) suggesting that the physical presence of the visitor was important for the interaction. This is similar to the results of the study by Pan and Steed (2016) where they found that the physical presence of a robot was better for communication, than a video or a virtual avatar. The local participants strongly expressed the feeling of being together with their friends during the experience, with comments such as, "*I was able to talk and feel the very strong presence of my friend. I could talk with her as if she was there with me and the movements of the robot reminded me of her gestures. From the movement of the head, I was able to understand what my friend was seeing.*" This statement particularly emphasizes the advantages of having a physical representation in the remote destination, instead of screen or desktop-based systems that do not allow people to express physical gestures and other non-verbal forms of communication. Furthermore, with a median score of 5, the question related to the effectiveness of the proxy system (*Interact_Local*) provides a satisfactory evaluation, since the response given by the locals was with respect to the entire experience, including the times when they were actually interacting with the proxy representation. Even though many locals did not immediately realize the moment the visitor had left their destination, they did eventually catch on to the fact that they were not present anymore. One participant said, "*I was talking with my friend about things we are currently working on together. At some point it was strange to notice that the context of the conversation changed out of the blue and I later realized that I was not really interacting with my friend anymore, rather with a pre-recorded audio of him.*" However, one local did not realize the change throughout the experience, and thought that the visitor was actually saying the statements that were being played back by the proxy. Their post-experience comments were, "*We had an informal conversation about Java classes, buying a dog, and holidays. Then my friend told me his favorite movie was The Gremlins and that he forgot something (… that he was saying).*" From the perspective of the visitors, one commented that they felt as if the locals had not realized that the visitor had left the destination: "*I could interact with 3 different people at one time and it felt like all the three thought I had not left the space I shared with them at all.*"

All participants were asked for overall comments about the system or suggestions for improvement, and most of the answers were related to the hardware limitations of the robots. Many expressed the desire to be able to walk in the destination. This feature had been implemented, but was disabled for the experiment to reduce the risks to the local participants and to limit the

chance of damage to equipment. Furthermore, the tasks had been designed specifically to not involve walking; however, the feature was still missed by some of the participants.

Body Ownership Illusion and Agency

Even though overall the visitors claimed to have felt a sense of body ownership and agency, there is a minor difference in the median scores between the two robots, likely due to the differences in physical dimensions and movement mechanisms. A participant mentioned, "*I did feel more like myself with the bigger robot, and less like myself with the smaller. I am unsure whether to attribute this only to the size of the robot, or not.*" Regardless of that, however, the positive responses for the question related to the BOI for both robots suggests that BOIs can be elicited regardless of robot size, given the first person perspective view over the robot body, head-based sensorimotor contingencies, and visuomotor synchrony for limb movement. This was shown earlier by van der Hoort et al. (2011) who were able to induce ownership over bodies of varying sizes. Although we were able to successfully elicit a BOI over the two robotic bodies, the median score for the variable *MyBody* is at the midpoint (Robothespian: median = 4, IQR = 2). This could be due to the fact that participants could not look all the way down to view their "robotic" body from 1PP (first person perspective) due to constraints in the movement of the neck of the robots, so that the main visual feedback they had with respect to "owning" the robot body was *via* the mirror placed in front of both robots at the destination. The only time when the visitors had a direct view of their robotic body was if their robotic arm would enter their FOV, while embodied in the robot. The evidence of this can be seen in the responses to the question which asks if the arm they saw was their own, with higher scores (*Arm_Mine: Robothespian: median = 5, IQR = 1; Nao: median = 5, IQR = 3*) than any other questions related to body ownership.

With respect to agency, high scores for the question *Agency* for both robots show that given the quite accurate mapping between their own and the corresponding robot movements, the robot sizes seem to not matter, a finding similar to the study by Banakou et al. (2013). However, some differences could arise depending on the hardware installed in the robots, with the Nao's control mechanism built with DC motors for all joints, whereas the Robothespian's arms having pneumatic motors. Although there are no differences in the responses to the question *ControlBody*, we suspect that the construction of the robot could influence the feeling of agency, since the electrical motors in the Nao with higher sensitivity may allow for a finer degree of control over the subtle arm movements, while the pneumatic motors of the Robothespian might be unable to represent these small movements made by the visitors and only tend to move once the movement is more noticeable. A follow-up study that focuses on the effect of the control mechanism of various humanoid robots on the feeling of agency could provide insight regarding this question.

These results are very promising and provide useful clues for future research in this field. Although the reported scores for the questions related to physical and psychological comfort were positive, the current setup can be a bit cumbersome, since the visitor is required to wear a full-body tracking suit that needs to be calibrated before they can be used, along with a heavy HMD that can cause discomfort. With companies such as Facebook, Microsoft, and Samsung now marketing virtual reality technologies, research and development of virtual reality hardware is gaining momentum, devices such as the Oculus Rift, HTC Vive, Microsoft HoloLens,[7] and Samsung Gear VR[8] are key to making virtual reality mainstream. Even though there have been advances in the field of markerless tracking technology with the Microsoft Kinect and Organic Motion,[9] this usually is at the cost of accuracy. Several tasks in the real world require a very high degree of precision, which might not be feasible with current tracking systems. On the other hand, humanoid robots available in the market also have a long way to go before they can be seriously used as remote surrogate bodies for such purposes. High cost of manufacture and constraints regarding hardware and safety issues need to be confronted in order to have a humanoid robot that can be used in a system such as this. In addition, it is also important to advance the current theoretical framework regarding BOIs with robots of different types, as it may shed light on the influencing factors, giving us clues as to which factors are more crucial than others. As observed in the experiment, people reported differences in feelings of not just ownership but also the way they felt the locals would perceive them in the destination. For example, a participant commented that they would not like to use the Nao for business meetings as they felt nobody would take them seriously, but would actually prefer the smaller robot for spending time with family. This shows the need for there to be humanoid robots that are physically appropriate for the tasks that they can carry out. Furthermore, for certain circumstances, it might be preferable to have a robot whose appearance could be customized according to the situation, as this would help overcome common issues of prejudice based on judgments made about physical or virtual appearance (Nowak and Rauh, 2008).

The other important result is that people were able to cope with being at three different places at the same time and were able to carry out tasks successfully. While numerous telepresence applications follow a one-to-one approach, where one operator controls one physical effector (robotic arm or humanoid robot) at the remote location, we have developed a system that allows the participant to be in three separate locations at the same time, and allows two-way interaction with the remote people at each location. Due to the proxy system, the interactions can continue even if the visitor is not present in the destination. Although we carried out the study with a simplified version of the proxy system, it can also be developed into a much more intelligent system, with a much higher degree of autonomy. The proxy could be task based, so it could have relevant behavior depending on the tasks that the visitor is required to do. It could also have the ability to "learn" new tasks according to the need, such as performing repetitive movements, or for example, learning a speech that the visitor needs to give. This could allow the visitor to do other activities

---

[7]https://www.microsoft.com/microsoft-hololens/en-us.

[8]http://www.samsung.com/global/galaxy/gear-vr/.

[9]http://www.organicmotion.com/.

as their surrogate robot could give a speech, and with a press of a button they could go to their destination and check on the proxy, in case it was needed. The proxy itself could contact the visitor in case they were unable to solve some issue on their own and needed the visitor's intervention.

We have introduced an instance of a possible type of system for multi-destination beaming, shown how the hardware and software can be put together, and shown that such a system might be useful in today's world of multiple simultaneous demands and meetings. However, there are many issues outstanding. For example, the type of tasks that could be accomplished, whether the proxy can reliably inform the visitor about what they have missed while away, detailed understanding of the relationship between robot capabilities and task performance, the role of body ownership and presence in facilitating performance, and not least ethical issues. On the latter point for the purposes of the experiment, participants acting as locals were not told in advance that their friend would not always be present but be substituted by a proxy, they were only debriefed afterward. This is suitable for an experimental study but may not be for real life applications. However, this situation is not particularly unusual today: when we are in a telephone call with someone, we do not know if their attention is really on a text message or email that they are simultaneously composing, or if they are reading a newspaper giving the occasional "Aha" and "OK" – essentially acting as their own proxy. These are anyway important issues beyond the scope of this paper.

As the hardware and software systems improve, the type of system presented above could become a viable solution for multiple simultaneous demands and could significantly reduce the need for traveling over long distances. The long-term vision is to have surrogate robot docking stations all over the world, where anyone could connect to whichever robot they wanted, and moreover as many robots as they wanted, and "teleport" to many locations instantaneously.

REFERENCES

Adalgeirsson, S. O., and Breazeal, C. (2010). "MeBot: a robotic platform for socially embodied telepresence," in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Osaka: IEEE Press), 15–22.

Banakou, D., Groten, R., and Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12846–12851. doi:10.1073/pnas.1306779110

Bankoski, J., Wilkins, P., and Xu, Y. (2011). "Technical overview of VP8, an open source video codec for the web," in *Proceedings – IEEE International Conference on Multimedia and Expo* (Barcelona: IEEE Computer Society), 1–6.

Barfield, W., and Hendrix, C. (1995). The effect of update rate on the sense of presence within virtual environments. *Virtual Real.* 1, 3–15. doi:10.1007/BF02009709

Bluethmann, W., Ambrose, R., Diftler, M., Askew, S., Huber, E., Goza, M., et al. (2003). Robonaut: a robot designed to work with humans in space. *Auton. Robots* 14, 179–197. doi:10.1023/A:1022231703061

Friedman, D., and Hasler, B. S. (2016). "The BEAMING proxy: towards virtual clones for communication," in *Human Computer Confluence Transforming Human Experience Through Symbiotic Technologies*, 156–174.

Friedman, D., Pizarro, R., Or-Berkers, K., Neyret, S., Pan, X., and Slater, M. (2014). A method for generating an illusion of backwards time travel using immersive virtual reality – an exploratory study. *Front. Psychol.* 943:1–15. doi:10.3389/fpsyg.2014.00943

Friedman, D., Salomon, O., and Hasler, B. S. (2013). *Virtual Substitute Teacher: Introducing the Concept of a Classroom Proxy.* London: King's College London, 186.

Glas, D. F., Kanda, T., Ishiguro, H., and Hagita, N. (2008). "Simultaneous tele-operation of multiple social robots," in *Proceedings of the 3rd International Conference on Human Robot Interaction – HRI '08* (Amsterdam: ACM), 311–318.

González-Franco, M., Pérez-Marcos, D., Spanlang, B., and Slater, M. (2010). "The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment," in *Proceedings – IEEE Virtual Reality Conference (VR)* (Waltham, MA: IEEE), 111–114.

Gooskens, G. (2010). "Where am I?": The problem of bilocation in virtual environments. *Postgrad. J. Aesth.* 7, 13–24.

Guizzo, E. (2010). When my avatar went to work. *IEEE Spectrum* 47, 26–50. doi:10.1109/MSPEC.2010.5557512

Iamblichus, C., and Taylor, T. (1926). *Iamblichus' Life of Pythagoras.* London: Watkins.

Kilteni, K., Bergstrom, I., and Slater, M. (2013). Drumming in immersive virtual reality: the body shapes the way we play. *IEEE Trans. Vis. Comput. Graph* 19, 597–605. doi:10.1109/TVCG.2013.29

Kilteni, K., Maselli, A., Kording, K. P., and Slater, M. (2015). Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. *Front. Hum. Neurosci.* 9:141. doi:10.3389/fnhum.2015.00141

ETHICS STATEMENT

The study was approved by the Comissió Bioètica de la Universitat de Barcelona. Participants were given an information sheet describing the experiment and gave written informed consent.

AUTHOR CONTRIBUTIONS

MS conceived the project, which was implemented by SK with the help of XM and PB. The experiment was carried out by SK with the help of PB. The proxy was designed by DF and implemented by KO-B. SK wrote the paper with the help of MS, and was reviewed by all authors.

ACKNOWLEDGMENTS

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/frobt.2016.00065

VIDEO S1 | A video that demonstrates the working of the multi-destination system as described in the study.

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

Kishore, S., González-Franco, M., Hintemüller, C., Kapeller, C., Guger, C., Slater, M., et al. (2014). Comparison of SSVEP BCI and eye tracking for controlling a humanoid robot in a social environment. *Presence Teleop. Virt. Environ.* 23, 242–252. doi:10.1162/PRES_a_00192

Kishore, S., Navarro, X., Dominguez, E., de la Pena, N., and Slater, M. (2016). Beaming into the news: a system for and case study of tele-immersive journalism. *IEEE Comput. Graph. Appl.* PP, 1–1. doi:10.1109/MCG.2016.44

Lincoln, P., Welch, G., Nashel, A., State, A., Ilie, A., and Fuchs, H. (2011). Animatronic shader lamps avatars. *Virtual Real.* 15, 225–238. doi:10.1007/s10055-010-0175-5

Maister, L., Sebanz, N., Knoblich, G., and Tsakiris, M. (2013). Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition* 128, 170–178. doi:10.1016/j.cognition.2013.04.002

Marescaux, J., Leroy, J., Gagner, M., and Rubino, F. (2001). Transatlantic robot-assisted telesurgery. *Nature* 168, 873–874. doi:10.1038/35096636

Maselli, A., and Slater, M. (2013). The building blocks of the full body ownership illusion. *Front. Hum. Neurosci.* 7:83. doi:10.3389/fnhum.2013.00083

Nagendran, A., Steed, A., Kelly, B., and Pan, Y. (2015). Symmetric telepresence using robotic humanoid surrogates. *Comput. Animat. Virtual Worlds* 26, 271–280. doi:10.1002/cav.1638

Normand, J. M., Giannopoulos, E., Spanlang, B., and Slater, M. (2011). Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. *PLoS ONE* 6:e16128. doi:10.1371/journal.pone.0016128

Normand, J. M., Sanchez-Vives, M. V., Waechter, C., Giannopoulos, E., Grosswindhager, B., Spanlang, B., et al. (2012a). Beaming into the rat world: enabling real-time interaction between rat and human each at their own scale. *PLoS ONE* 7:e48331. doi:10.1371/journal.pone.0048331

Normand, J. M., Spanlang, B., Tecchia, F., Carrozzino, M., Swapp, D., and Slater, M. (2012b). Full body acting rehearsal in a networked virtual environment – a case study. *Presence* 21, 229–243. doi:10.1162/PRES_a_00089

Nowak, K. L., and Rauh, C. (2008). Choose your "buddy icon" carefully: the influence of avatar androgyny, anthropomorphism and credibility in online interactions. *Comput. Human Behav.* 24, 1473–1493. doi:10.1016/j.chb.2007.05.005

Pan, Y., and Steed, A. (2016). A comparison of avatar-, video-, and robot-mediated interaction on users' trust in expertise. *Front. Robot. AI* 3:1–12. doi:10.3389/frobt.2016.00012

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

Perez-Marcos, D., Solazzi, M., Steptoe, W., Oyekoya, O., Frisoli, A., Weyrich, T., et al. (2012). A fully immersive set-up for remote interaction and neurorehabilitation based on virtual body ownership. *Front. Neurol.* 3:110. doi:10.3389/fneur.2012.00110

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

Preston, C., Kuper-Smith, B. J., and Ehrsson, H. H. (2015). Owning the body in the mirror: the effect of visual perspective and mirror view on the full-body illusion. *Sci. Rep.* 5, 18345. doi:10.1038/srep18345

Raskar, R., Welch, G., Low, K., and Bandyopadhyay, D. (2001). "Shader lamps: animating real objects with image-based illumination," in *Proceedings of the 12th Eurographics Workshop on Rendering Techniques* (Vienna: Springer), 89–102.

Sanchez-Vives, M. V., and Slater, M. (2005). From presence to consciousness through virtual reality. *Nat. Rev. Neurosci.* 6, 332–339. doi:10.1038/nrn1651

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3549–3557. doi:10.1098/rstb.2009.0138

Slater, M., and Sanchez-Vives, M. V. (2014). Transcending the self in immersive virtual reality. *Computer* 47, 24–30. doi:10.1109/MC.2014.198

Spanlang, B., Navarro, X., Normand, J.-M., Kishore, S., Pizarro, R., and Slater, M. (2013). "Real time whole body motion mapping for avatars and robots," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology – VRST '13* (New York, NY: ACM Press), 175.

Steed, A., Steptoe, W., Oyekoya, W., Pece, F., Weyrich, T., Kautz, J., et al. (2012). Beaming: an asymmetric telepresence system. *IEEE Comput. Graph. Appl.* 32, 10–17. doi:10.1109/MCG.2012.110

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

Tsui, K. M., Desai, M., Yanco, H. A., and Uhlik, C. (2011). "Exploring use cases for telepresence robots," in *6th ACM IEEE International Conference on Human Robot Interaction* (New York, NY: ACM), 11–18.

van der Hoort, B., Guterstam, A., and Ehrsson, H. H. (2011). Being Barbie: the size of one's own body determines the perceived size of the world. *PLoS ONE* 6:e20195. doi:10.1371/journal.pone.0020195

Wissmath, B., Weibel, D., Schmutz, J., and Mast, F. W. (2011). Being present in more than one place at a time? Patterns of mental self-localization. *Conscious. Cogn.* 20, 1808–1815. doi:10.1016/j.concog.2011.05.008