OPEN

Substitutional Reality System: A Novel Experimental Platform for Experiencing Alternative Reality

Keisuke Suzuki, Sohei Wakisaka & Naotaka Fujii

Laboratory for Adaptive Intelligence, Brain Science Institute, RIKEN 2-1 Hirosawa, Wako, Saitama, Japan.

We have developed a novel experimental platform, referred to as a substitutional reality (SR) system, for studying the conviction of the perception of live reality and related metacognitive functions. The SR system was designed to manipulate people's reality by allowing them to experience live scenes (in which they were physically present) and recorded scenes (which were recorded and edited in advance) in an alternating manner without noticing a reality gap. All of the naïve participants ($n = 21$) successfully believed that they had experienced live scenes when recorded scenes had been presented. Additional psychophysical experiments suggest the depth of visual objects does not affect the perceptual discriminability between scenes, and the scene switch during head movement enhance substitutional performance. The SR system, with its reality manipulation, is a novel and affordable method for studying metacognitive functions and psychiatric disorders.

Have you ever thought that what you were experiencing could be a dream or that friends you were talking to would disappear when you blinked? In principle, we believe what we see, but is it the case that what we see is necessarily really happening? A more accurate statement could be that we see what we believe. Consciously or unconsciously, we have a strong conviction that we experience live, ongoing reality. We refer to this simply as having a conviction about reality (CR). Usually, CR is falsely maintained in dreams. Consider the movie "Inception", in which people were unable to discriminate between reality and dreams. To return to reality, they needed a physical "kick", or a clue prepared as an emergency key. What happens if we do not have the clue once we are trapped in the dream? This type of disorientation is not limited to science fiction; similar occurrences are a part of some psychiatric diseases[1–6].

During periods when we are awake, we usually do not need such an explicit clue because the maintenance of a CR is a basic metacognitive function that humans have ("cognition of cognition"). Although the definition of metacognition has not been fully established, introspection, confidence and self-monitoring are also considered metacognitive processes that relate to each other[7–12]. Clinical studies have shown that CR is a key issue for understanding metacognition. For example, disoriented patients cannot properly recognise time, objects or people in reality[1]. These patients often confabulate their ongoing reality, creating stories that are clearly inconsistent with their current situation (e.g., reduplicative paramnesia, geographical mislocation, and spontaneous confabulation)[2–6]. These confabulations are a result of metacognitive dysfunction in that these patients seem to lose the appropriate introspections to their cognitions.

Recent psychological studies examining "choice blindness" have revealed that confabulation with regard to reality can be induced in normal healthy participants by manipulating the outcome of their decisions using a simple sleight of hand (e.g., exchanging cards or a trick jam container)[13–15]. In these experiments, participants selected a card and were then asked to justify their decision, either with or without the card being switched. A significant number of participants did not notice the switch and proceeded to confabulate reasons for selecting the card that they did not in fact select, apparently violating introspective consistency. However, if their CR was weakened (i.e., when they started doubting that reality was manipulated and thus not as they subjectively experienced, in this case, by becoming aware of the sleight of hand), the frequency of such confabulations drastically decreased. If the experimenter explained the trick, none of the participants confabulated because their CR had disappeared. In another study, when a virtual agent presented the card trick on a computer screen, people noticed the trick easily[16]. These studies suggest that 1) reality manipulation is a promising tool to investigate metacognitive function and 2) CR should be maintained for the manipulation to be successful.

In this report, we describe an experimental setup that allows novel types of reality manipulation while maintaining participants' CR, substantially extending previous reality manipulations utilized in cognitive science such as the choice blindness studies described above. In this setup, participants' live reality was covertly substituted with an alternative reality without their noticing the change; thus, their CR remained intact. This situation is referred to as substitutional reality (SR) and our implementation of SR as "SR system", in which participants can experience live scenes and previously recorded scenes as equally realistic such that everything in these scenes seems to exist in the surrounding physical reality. The SR system implements and extends several techniques that have been used in virtual or mixed reality (VR or MR) systems (a head-mounted display (HMD) and a panoramic video camera). VR/MR systems have been broadly and successfully used in psychology, cognitive neuroscience, and various therapies[17]. We will describe the SR system configuration in the next section, as well as discuss its advantages and disadvantages with respect to VR/MR systems in a later discussion section.

However to introduce the SR system, we first consider an example of SR-based reality manipulation *with CR maintained*, that is easily achievable by the SR system, but would be technically very difficult or, in some cases, impossible with any other methods, including VR/MR systems. In our example, we can present a realistic experimental room with experimenters working to set something up or even speaking to the subject, without the subject noticing that the entire scenario is in fact not happening. Additionally, we can cause participants to experience inconsistent or contradictory episodes, such as encountering themselves. Another example is experiencing identical episodes repeatedly (e.g., conversations or one-time-only events, such as breaking a unique piece of art). Such episodes create a déjà vu-like rare situation in that participants experience the same event repeatedly in their live reality, and they are sure that the same event happened before. Visual experience of the world with different natural laws (i.e., weaker gravity or faster time) can also be implemented. If we consciously experience these events and yet believe them to be real, how do we perceive/recognise them? How does our brain manage the inconsistencies? Do we deceive ourselves with confabulations

or somehow discover the substitutions and lose a CR? Even if a CR is maintained in these episodes, we may experience an uncertainty about the reality of the situation. How is this uncertainty manifested, both behaviourally and in terms of physiological signals? Using the SR system, these important questions can be investigated, allowing the SR system to be a novel and affordable method for studying metacognitive functions.

Results

Implementation of the SR system. The SR system consists of the following three sub-modules: a recording module, an experience module and a control computer. The recording module (Fig. 1, left) was equipped with a microphone and a panoramic video camera with the ability to record a panoramic movie, which was then stored on the control computer. The experience module (Fig. 1, right) consisted of a HMD, a head-mounted camera, an orientation sensor, noise-cancelling headphones and the same microphone used by the recording module. The camera was mounted at the front of the HMD, and the orientation sensor was mounted on a rim. The experience module alternately presented two different types of scenes: the first was a real-time scene captured by the head-mounted camera and the microphone (live scene) and the second was a scene that was previously recorded and edited in advance by the recording module (recorded scene). During presentation of the recorded scene, the panoramic movie was cropped in real-time to fit the HMD display size. The cropped area was determined based on the participant's head orientation, which was obtained from the orientation sensor (i.e., when a participant turned to the left, the cropped area shifted accordingly). Therefore, assuming that the head was kept stable in a position, natural visuo-motor coupling was ensured both in the live and recorded scenes. Additionally, by setting the head position close to the location where the panoramic camera was placed when recording the movie, the visuo-motor experiences of live and recorded scenes were similar enough to be indistinguishable. In both scenes, an identical image was presented to each eye, meaning that there was no binocular parallax. In this way, participants' reality could be manipulated by
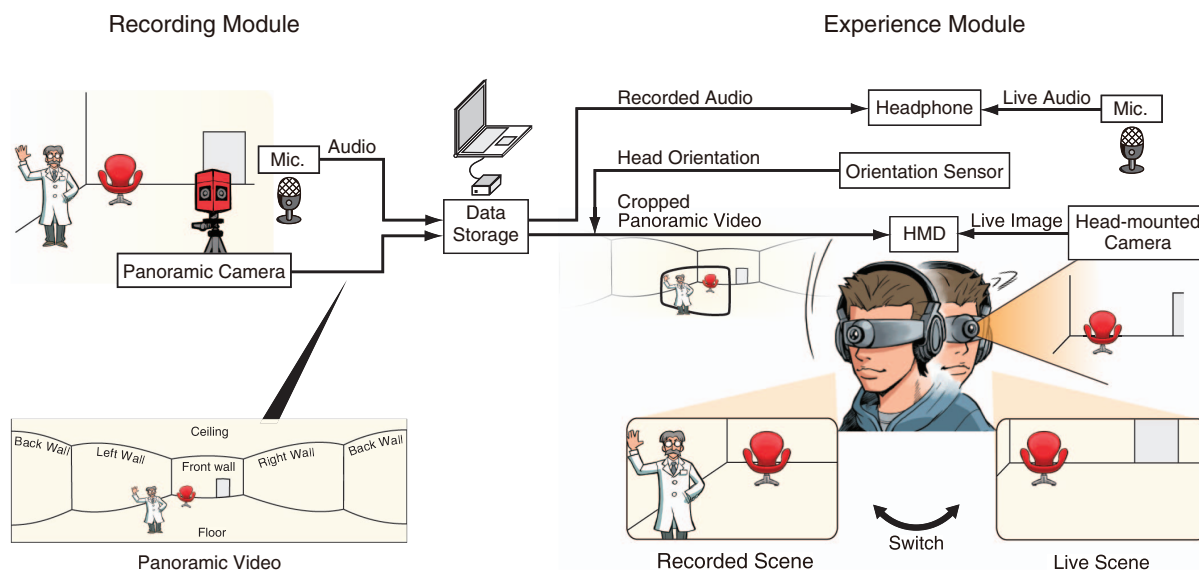
Figure 1 | Substitutional Reality System. In the recording module (left), the panoramic view was recorded in advance by a panoramic camera, and stored in the data storage connected to the control computer. In the experience module (right), either a live scene captured by a head-mounted camera or recorded scenes cropped from a pre-recorded movie were shown on a head-mounted display (HMD). The cropped area presented in the recorded scenes was determined in real-time using head orientation information calculated from the HMD orientation sensor. Scene examples are shown here. In the recorded scene a person with a lab coat waved his hand, who was not present in the live scene. A participant believed the person with the lab coat was physically present there, when the covert switch from the live to the recorded scene was successfully performed.

Figure 2 | A cartoon depiction for each step of Experiment I's sequence is shown. (a) During the recording session, the participant was invited into the room and received instructions about the experiment. During this time, everything was recorded for the Doppelgänger scene. (b) Normal Question scene. After the covert substitution from the live scene to the recorded scene, the participant replied naturally to the experimenter, indicating that the substitution was successful. (c) Doppelgänger scene. The participant saw himself, thereby realising that the scene he had experienced was not live. (d) Fake Live scene. The SR system worked even after the Doppelgänger scene. Seven of 10 participants could not detect that the given scene was recorded. (e) The Live scene after the Fake Live scene. The participant was not certain whether he was experiencing live or recorded scenes any more. See DISCUSSION. Colour bars at the right of each box indicate scene differentiation (orange for a live scene and green for a recorded scene). For convenience, the microphone and connection cables are omitted from the drawings.

covertly switching the live scene and the recorded scenes back and forth.

The experience of the SR system was determined by the scene sequence (including the live scene), which could be either fixed (as in following Experiment I), or manually adapted by experimenters depending on the response of participants. Such manual sequence manipulation is feasible when more complex and interactive scene selection is required.

Performance of the SR system (Experiment I). We assessed the performance of the SR system ($n = 21$, see Methods regarding Experiment I) by observing the following three points: (1) whether the SR system could covertly substitute reality successfully, (2) how a participant's CR was modulated when exposed to an unrealistic, extremely contradictive event and (3) whether we can re-establish participants' CR after they explicitly noticed the substitution and mechanism of the SR system.

To address these questions, we designed a sequence of scene presentations. A five-frame comic strip depicts how the sequence was presented (Fig. 2). We employed three scenes that were recorded prior to the experience session. Each scene corresponds to each of three questions described above, respectively. The first scene was designated the "Normal Question" scene, in which the experimenter appeared and asked several questions (e.g., "Do you feel OK with HMD?" or "Can you look around?"). In this case, the experimenter was pretending to speak to the participant during the recording session, although the experimenter was actually speaking to the panoramic camera. The second scene was extremely contradictive and referred to as a "Doppelgänger" scene, in which the participant appeared from the door with the experimenter, walked close to the panoramic camera, had a conversation with the experimenter (2~3 minutes) and walked out of the room. This scene was recorded when the participant was invited into the experimental room to receive instructions (Fig. 2a). The third scene was a "Fake Live" scene, in which the experimenter behaved as if he was talking in real-time, saying, "So, this is the live scene. I'm here. Can you tell?" (Fig. 2d).

During the experiment, we instructed the participants to sit back in the chair with their hands resting on their thighs and to freely look around the room, but not to look down at themselves because their body would not be visible in the recorded scenes. Each participant first experienced a live scene via the head-mounted camera and the microphone. During the live portion of the experiment, the experimenter asked questions that were similar to the ones asked in the Normal Question scene and confirmed that the HMD was comfortable. When the participant moved his/her head, the experimenter manually switched the live scene to the Normal Question scene. Switching during head movement enhanced substitution performance. This issue is described in Experiment III. If the participant did not look around the room spontaneously, we asked him/her to do so.

During the experiment, all of the participants verbally responded to the experimenter's questions in the Normal Question scene as if the scene were taking place in real time (Fig. 2b; additionally, see the supplementary video S1). Afterwards, all of the participants reported that they did not notice the switch and that they believed they were experiencing actual events throughout the entire session. This result shows that (1) without any prior knowledge about SR system, people did not recognise the substitution and (2) an interaction could be established with people appearing in previously recorded scenes (in this case, a fake conversation including simple questions and responses).

Next, we switched the scene to the Doppelgänger scene (Fig. 2c and the supplementary video S1 at 1 : 32). When the participants saw themselves in the recorded scene, all of the participants became aware that they were not experiencing live reality. Not surprisingly, the Doppelgänger scene was too contradictory to maintain a CR.

Finally, we switched the scene to the Fake Live scene (Fig. 2d and the supplementary video S1 at 2:07). Ten of the 21 participants experienced this optional scene after the Doppelgänger scene. Seven of them could not detect that the given scene was the recorded scene. We confirmed this from their replies to the experimenter in the scene (e.g., "Yes, I know this is live, of course"), indicating that they re-established CR. The remaining three noticed that the scene had been recorded previously, stating that they noticed a difference in the sound quality between the live and the Doppelgänger scene and used this auditory difference as a cue in the Fake live scene. At the end of Experiment I, we switched back to the live scene and explained that the previous Fake Live scene was also a recorded scene. The participants who did not detect the substitution during the Fake Live scene were often confused during this conversation because their conviction became uncertain (Fig. 2e and the supplementary video S1 at 2 : 45).

We observed an interesting behaviour in one participant during the Normal Question scene. The participant happened to raise his hand in front of his eyes, although he had been instructed not to do so. Although his hand was invisible to him, he did not notice the switch and continued to respond to the experimenter's (recorded) questions. After the experiment, he reported that he was confused when he could not see his hand, but he thought that he might have put his hand somewhere other than in his field of view. Although an "invisible hand" would seem to be strongly contradictory, the reality substitution worked, and the contradiction was compensated for with confabulation. This observation suggests that participants' CR can be maintained even in apparently contradictive situations with strong conviction.

In the following studies, we designed two verification experiments to manipulate two important major factors (i.e., motion parallax and the scene switch timing) to determine how they influenced substitution performance.

Difference in motion parallax (Experiment II). When head position changes, the shape and depth of objects in the visual field change accordingly. Even when head position is stationary, changing the orientation of the head can alter the shape of objects (motion parallax). Although there was normal motion parallax in the live scenes, it was absent in the recorded scenes in the SR system because the viewpoint of the panoramic camera was fixed. Therefore, if the participants paid attention to the difference in motion parallax when changing their head position or orientation, they would be able to differentiate live and recorded scenes. However, it has to be emphasised that none of the participants in Experiment I spontaneously noticed the difference in motion parallax, even after they were informed about the substitution trick. They kept looking around at visual objects at various depth (~1.5 m), but could not use the parallax difference as a clue until we explained it. This suggests that the visuo-motor experience could be

natural enough without motion parallax in the SR system and that object distance may play a minor role in influencing successful substitution.

To examine this proposal, in Experiment II we tested the effect of motion parallax on substitution performance when it was explicitly explained and used by participants as a discrimination clue. The participants ($n = 10$) were told about the mechanism of the SR system, then asked to sit alone in a room, where one red chair was placed in front of them (Fig. 3a–c). Each participant was asked to determine whether the scene he/she was viewing was live or recorded by monitoring the motion parallax around the red chair that was induced by his/her own head motion. There were three different distances (1.0 m/2.5 m/4.0 m) between the participant and the chair (Fig. 3a). In general, longer distances cause less motion parallax. To introduce the wide variety of head movements, participants received two instructions with a randomised order (Fig. 3b). With "Head Only" instructions, the participants were asked to rotate their head without body displacement. With "Head and Upper-body" instructions, the participants were asked to displace their upper body (i.e., move their shoulders) and change their head orientation to induce greater motion parallax. Figure 4a shows the correct detection rates for each distance. As we expected, the correct detection rate was higher in "Head and Upper-body" instruction than in "Head Only" instruction. But a statistical comparison did not show significant differences between the three distance conditions [Friedman test: $p = 0.627$] in both instructions. Figure 4b shows the time lag between scene switching and correct detection in the six conditions. A two-way repeated measures ANOVA revealed a significant main effect of distance [$F(2,18) = 4.85, p < 0.05$], with no significant main effect of displacement [$F(1,18) = 3.37, p > 0.05$]. Multiple comparisons showed a significant effect between the 1.0 m and 4.0 m conditions (Scheffé's test: $p < 0.01$). There was no significant distance-by-displacement interaction ($F = 0.0648, p = 0.94$). Although the motion parallax is an important factor for the SR system performance, the high and constant correct rates regardless of the different distances indicates that the object distance does not necessarily affect the subjective discriminability of scenes. This finding is consistent with the observation in Experiment I that participants did not spontaneously find the difference in motion parallax, even though they looked around at objects that had different distances. It is important to note that we need to further investigate applying different environments in the SR system to generalise the results.

Head speed and detection rate of scene switching (Experiment III). Although head orientation was the same, the images from the live and recorded scenes could not be identical due to fluctuations in the orientation sensor and motion parallax. Thus, the image inevitably slipped at the switch onset between the live and recorded scenes. In Experiment I, to prevent the participants from noticing the visual slip, we heuristically switched the scenes manually only when the participants moved their heads so that the slip was perceptually masked during the scene transition. Although this worked well, it did not provide an appropriate range of head speeds for successful substitution. Here, we attempted to determine the optimal range of head speeds for successful switching in the SR system.

In Experiment III (Fig. 5), the participants were instructed to sit in a chair, to make their head position stable according to the Head Only instruction from Experiment II, and to look at different orientations by turning their head intermittently at one of four speeds: "Motionless" (<32 deg/sec), "Slow" (32–64 deg/sec), "Fast" (64–96 deg/sec) and "Very Fast" (>96 deg /sec) (Fig. 5a). The speed of the "Very Fast" condition roughly corresponds to the speed attained when an individual turns around quickly. Head speed was monitored by the orientation sensor on the HMD, and scene switching occurred when the speed exceeded the given instructed speed (see Fig. 5b).
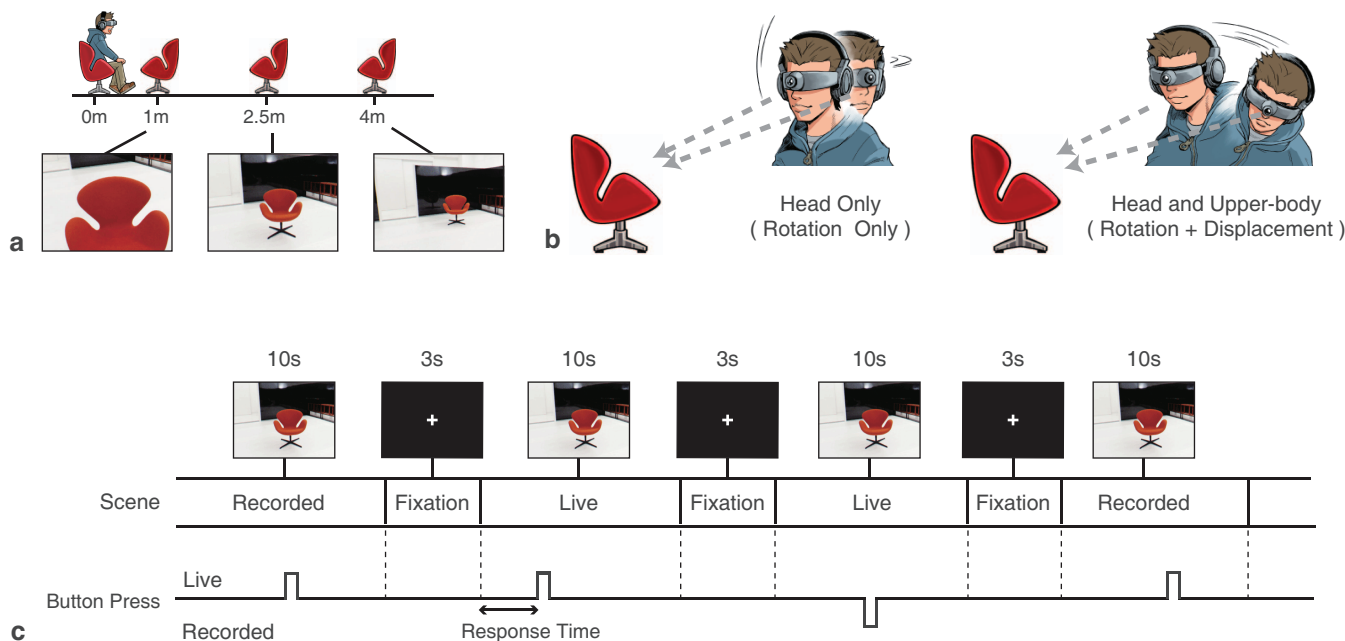
Figure 3 | Experimental Design of Experiment II. Two independent conditions were applied. (a) In the first condition, there were three different distances (1 m, 2.5 m, and 4 m) from an object in the visual field, presumably providing different degrees of motion parallax. (b) In the second condition, there were two different instructions for head movement. With the "Head Only" instruction, the participants could only change their head orientation. With the "Head and Upper-body" instruction, the participants could move their upper body in addition to their head. In both cases, the participants were instructed to keep their eyes on the chair (the line of sight is indicated with a grey dashed arrow). (c) Temporal sequence of Exp. II for discriminating between live and recorded scenes. Live scenes or recorded scenes were pseudo-randomly selected and presented (10 sec) interspersed with a 3 sec fixation period. The participants were asked to report whether the scene was live or recorded by pressing a button.

Participants were asked to focus on the onset of the scene switch and press a button on an interface box as soon as they detected the switch. Figure 6 shows the correct detection rates for the four speed conditions. A one-way repeated measures ANOVA revealed a significant main effect of speed ($F(3,27) = 19.38$, $p < 0.01$) (Fig. 6). Multiple comparisons showed significant effects between the "Motionless" condition ($76\pm2\%$) and the other three conditions ($45\pm3\%$, $36\pm2\%$, and $21\pm2\%$ for "Slow," "Fast," and "Very Fast", respectively) (Scheffé's test: $p < 0.001$), indicating that switch detection was easier when the participants did not move their head, with even "Slow" head motion significantly reducing the detection performance. Detection performance of visual changes decreases during head movements with HMD (i.e., head movement suppression[18]). The result suggests that the same suppression also occurred in our system, which hid the visual slip during the scene switch.

Discussion

The SR system, our novel video camera-based implementation of SR, allowed participants to experience recorded scenes subjectively as live scenes even prevented participants from doubting their FALSE perception. Importantly, the SR system is highly flexible in that a large repertoire of pre-recorded scenes can be used, so long as they can be recorded by panoramic camera and edited in advance.
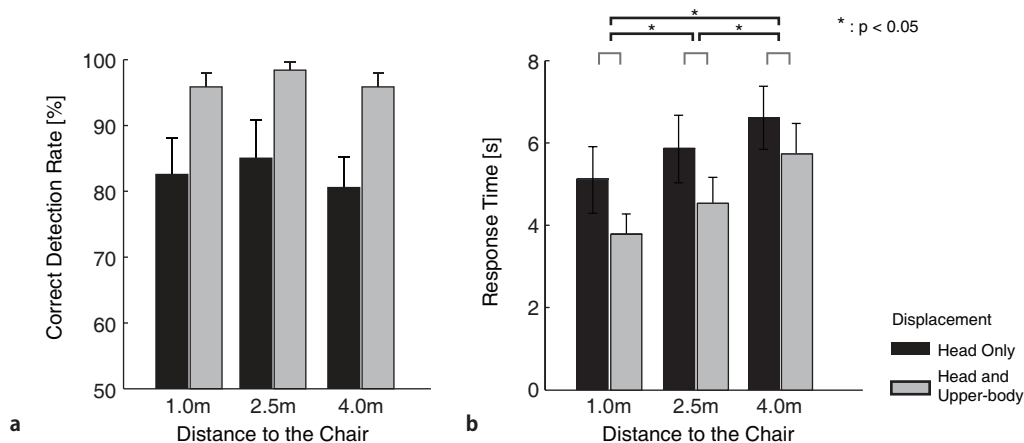
Figure 4 | Results of Experiment II. (a) Correct detection rates for the three distance conditions in Experiment II are indicated. All data were averaged across the participants ($n=10$). No significant difference was observed (Friedman's test: $p = 0.627$) between conditions. (b) Response latencies for the six conditions are shown. A two-way repeated measures ANOVA revealed a significant difference between the distance conditions. The p-values were obtained through post-hoc analysis (Scheffé's test). Error bars indicate the mean $\pm$ standard error. * indicates significance levels ($p < 0.05$).
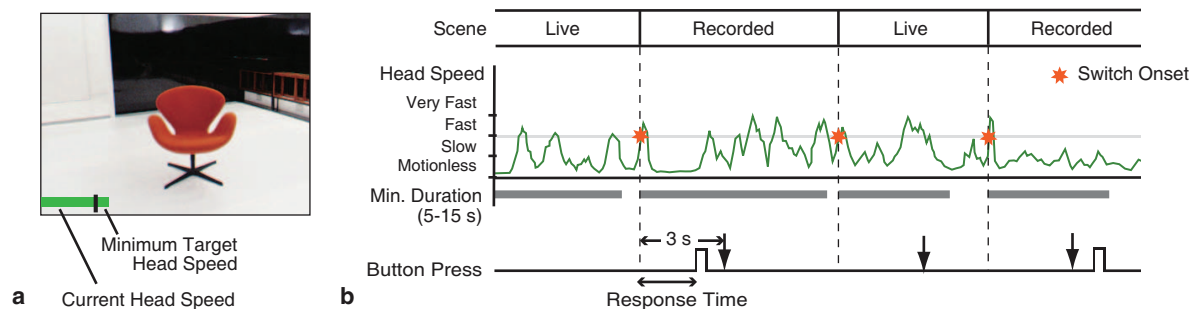
Figure 5 | Experimental Design of Experiment III. (a) Screenshot of the HMD. A bar at the bottom of the display indicated the current head speed, and a vertical line indicated the target head speed threshold. When the head speed exceeded the threshold, the bar changed from red to green. The participants were asked to maintain the target head speed. Scene switches occurred only when the participant's head speed exceeded the target speed (bar in green). Participants were asked to press a button upon identifying a scene switch. (b) Example time course of Experiment III for detecting the switch with the "Fast" instruction. The green line indicated actual head speed. A switch occurred after a short time had passed (randomly chosen from 5 to 15 sec) and when the head speed exceeded the instructed head speed. A response (button press) within 3 sec of the switch occurring was considered a correct response.

We showed (Experiment II/III) that a major factor influencing successful substitution in the SR system was consistent visuo-motor coupling throughout the experience. Due to this coupling, the participants could observe the environment naturally in both realities. When participants were engaged in the SR system, the experience always started with the live scene, although visual and auditory stimuli were provided indirectly via the HMD and headphones. This process induced strong CR in the SR system. Once conviction was established, it persisted even after the recorded scenes replaced the live scenes. In other words, the participants subjectively experienced the recorded scenes as being live reality.

Does the SR system become useless when participants notice its mechanism? The answer is no, given that even after they detected the substitution by experiencing events that contradicted reality or were debriefed about the mechanism by the experimenters, the majority of the participants (70%) could not detect the "Fake Live" scene (Experiment. I) indicating they re-established CR. Additionally, they often confused even when they later experienced the live scene with live conversation. The detailed analysis

of the reality confusion during live scene experience (i.e., what aspect of reality they began to question) remains for future investigation. The remaining participants detected the substitution, not by the visual slip, but due to a subtle difference in the auditory stimuli. Therefore, an improvement in auditory management may improve the substitution performance.

The characteristic feature of the SR system is the ability to manipulate the participants' subjective reality in ways that no other method can. However, for successful substitution, two major factors must be carefully managed. One is motion parallax, which only exists in the live scene. We confirmed that discrimination performance was not significantly affected by the location of the visual objects in the scenes, although motion parallax, when participants attended to it, functioned as a discrimination clue between scenes (Experiment II). Previous studies of depth perception with HMD suggested that motion parallax is not a major determinant in judging the depth and the size of visual objects[19], which might explain why the location of objects did not affect the discrimination performance.

The second factor was the visual slip that occurred during the scene switch. We found that the detection rate of scene switching could be significantly suppressed by enacting the switch when participants moved their heads (Experiment III), even at slow speeds. The result is also consistent with previous findings that the perceptual performance (sensitivity for the stimulus change, etc) were suppressed during the head movement (head movement suppression)[18,20–22]. Such suppression has been already incorporated into the VR technique (e.g., redirected walking[23]). The scene switch between live and recorded scenes during head movement can be considered as another application of the head movement suppression.

Besides careful management of these factors, there are still several practical concerns that have to be solved for introducing the SR system. For instance, invisible self-body in the recorded scene is one of the biggest concerns in the SR system. We minimized the impact of the concern by asking the participants not to look down their body during the experiment. However, it is not easy when the experiment lasts longer. One possible solution is physically covering participants' hands and lower body. Another solution is using Chroma keying and extracting participants' body image from HMD camera stream and overlay the image on the recorded scene. This is technically possible and may solve the problem. Another concern is a budget issue since a commercial panoramic video camera is expensive (the initial cost for setting up SR system is about $30K. To implement more affordable system, one option is to employ a combination of a digital camera and the one-shot
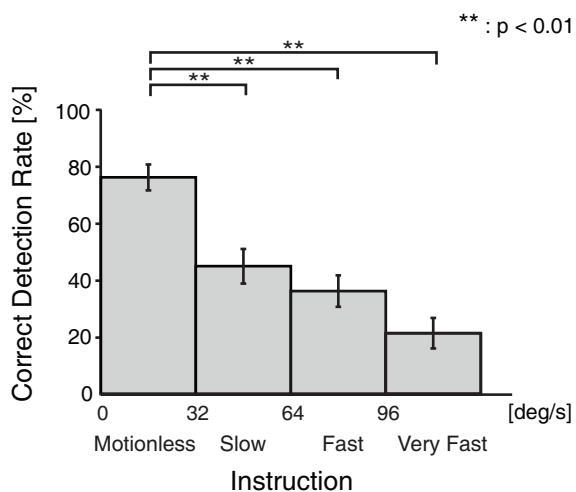
Figure 6 | Results of Experiment III. All data were averaged across the participants (n = 10). (a) Correct detection rates for the different head speeds at switches. A one-way repeated measures ANOVA revealed a significant difference between the instructions. The p-values were obtained through post-hoc analysis (Scheffé's test). Error bars indicate the mean ± standard error. ** indicates significance levels ($p < 0.01$).

panoramic lens mirror. The system might be able to substitute the reality to some extent, but there will be more limitations than the current system because of lower visual quality and narrower recordable angle range.

What is the difference between the SR system and a conventional immersive VR system? Current VR technologies with highly realistic computer graphic (CG; i.e., using the texture/video-texture mapping technique), high screen resolutions, fast frame rates and other VR specifications provide a strong feeling of presence (i.e., the feeling of "being there")[24–26]. Additionally, the VR environment can be implemented such that participants can move freely within the environment, look at their own virtual bodies and touch visible objects. Such environmental interactions are crucial factors for enhancing the feeling of presence[27,28]. Importantly, when the contents of the experience in the virtual environment are plausible, the participant tends to react as if the contents are real, even if he/she is fully aware that they are not real[29]. As noted, due to the strong feeling of presence and the flexibility in constructing the virtual environment, virtual environments have been widely used in broad areas related to cognitive science[17]. However, our primary concern in this report is a CR, which is apparently similar to but still different from the feeling of presence by definition. Although the SR system has restrictions regarding environmental interactions (e.g., a participant cannot move around within it), the SR system can make participants feel that the events, people and anything in the recorded scenes physically exist in front of them.

Is there any other method that can implement reality substitution other than the SR system? Indeed, there are several other known technologies available for substitution. For example, the mixed reality (MR) system[30] and its variation, diminished reality (DR) system[31] overlaid computer graphics (CG) in the real scene presented through HMD. These systems can substitute reality if participants do not notice the reality gap with CG and the real scene.

The MR and the DR system allow more environmental interactions than the SR system allows (i.e., participants can move more freely). However, compared to these technologies, SR system is easy to use for daily operation; no need to struggle with filling in the reality gap with CG, therefore neither a computer graphics engineer nor VR studio is necessary.

"Winscape" (http://www.rationalcraft.com/Winscape.html) can be considered as another implementation of SR, as it can convince participants that a flat monitor on a wall, which shows a video stream of distant landscape, is a real window. The realistic feeling is enhanced by a "head-coupled perspective" that changes the display image based on displacement of the observer's viewing point while maintaining the proper perspective[32]. With Winscape, participants do not have to wear an HMD. However, the substitution can be made only through the display window. In conclusion, with regard to reality substitution, each of these technologies, including SR system, has advantages and disadvantages. We can choose one of them or combine them, depending on what type of reality substitutions are needed.

The combination of a panoramic camera and HMD with an orientation sensor has been used in previous studies[33,34]. These studies have mostly endeavoured to study telepresence, in which people experience scenes from a distant location. Technically, SR system can be considered an implementation of a novel variation of telepresence, which covertly shifts time without changing location, although this idea has never been implemented.

The SR system is widely applicable to experiments in which a CR needs to be maintained or manipulated. In particular, this system provides a novel tool for studying how metacognitive functions are affected when reality is manipulated in various, sometimes abnormal, ways. The Doppelgänger scene was found to be too contradictory (Experiment I), and all participants immediately lost their CR when they saw their own image. However, this is an extreme example. Rather, we can introduce moderate contradictions that do not negatively impact a participant's CR, yet may introduce uncertainty about ongoing events (see examples of substitution described in the introduction section). In other words, the SR system can surreptitiously introduce the mismatch between expectation and experience (i.e., prediction error[35]), which may be an important factor in the delusion formation not only by normal healthy people[36] but by psychiatric patients[1–6,37–39]. We expect that the analysis of participants' response (some of them are indeed expected to be delusive; e.g., delusive mislocation of the 'invisible hand' observed in Experiment I), contributes to better understanding of mechanism of delusion. For the comparison with SR-induced delusion, it may be necessary to make delusive patients also experience with the SR system. To do so, careful establishment of ethical procedures are required.

VR technologies have already been accepted as useful tools for psychological therapy for posttraumatic stress disorder (PTSD) and other types of phobias[40–42]. In this type of therapy, patients experience replicated episodes that are related to their trauma or phobia through immersive VR equipment. It is known that repeated exposure to traumatic episodes in a VR system often decreases the level and frequency of a particular trauma or phobia. The therapy's success is dependent on the feeling of presence in a VR system[17,25]. Thus, the following question naturally arises: what is the therapeutic effect of episodes that are provided with a CR? In other words, what if a given episode is real, not 'as if it is real', as in previous therapies? Although the effect remains unknown and appropriate ethical procedures should be established in future investigation, we expect that a CR with the SR system will add new directions to psychological therapy.

Our SR system is a novel method that allows the manipulation of reality and uncertainty in normal participants. These manipulations can serve as useful tools for understanding the mechanisms of metacognitive functions and psychiatric diseases. Additionally, this system has the potential to be a useful communication and entertainment platform given its outstanding substitutional performance in reality management.

Methods

All experimental procedures were approved by the RIKEN ethical committee [approval no. Wako 3rd, 20-4(4)]. All of the participants provided informed consent prior to the experiments.

Configuration of the SR System. The recording module (Fig. 1, left) consisted of a panoramic video camera (Ladybug3, Point Grey Research, BC, Canada) and a microphone (H2 Handy Recorder, ZOOM, Tokyo, Japan). The panoramic camera captured 6 movies in different orientations at 16 frames per second (fps) and combined them into a seamless panoramic movie (2048×1024 pixels). The area corresponding to a downward angle of 70–90 degrees below the horizon was not recordable with this camera and was left blank. The movie was stored on the data storage connected to the control computer (CPU: Core i7-940XM, Intel, California, US; GPU: GeForce GTX 260 M 1 GB, NVIDIA, California, US; OS: Windows7, Microsoft, Washington, US). The experience module (Fig. 1, right) consisted of an HMD (resolution: 640×480 pixel, VR920, VIZUX, New York, US), a CCD camera (CCD-V21, Sanwa Supply, Okayama, Japan, 16 fps), an orientation sensor (InertiaCube3, Intersense, Massachusetts, US), noise cancelling headphones (ATH-ANC7b, Audio-Technica, Tokyo, Japan) and the same microphone used in the recording module. To achieve a first-person perspective, the CCD camera was mounted at the front centre of the HMD (head-mounted camera), and the orientation sensor was mounted on the HMD rim. The visual properties of the two scenes (e.g., brightness, contrast) were matched by adjusting the properties to minimise clues that would allow live and recorded scenes to be discriminated. The latencies of visual feedback were within 100 msec in both scenes. The custom software (c++/openGL/openCV) managed whole operations. The keyboard connected to the computer was used for manipulation of scene sequence and switches by experimenters.

Experiment I: Performance of the SR system. *Participants.* Twenty-one adult volunteers (14 males and 7 females) served as paid participants (average age: 31.8 years). All had normal or corrected-to-normal vision. Two participants had previous experience in immersive virtual environments. Due to the nature of the experience, the participants were not informed about the mechanism of the SR system beforehand. They were asked to evaluate verbally the user experience of our newly

developed immersive human interface. They were informed that the experiment was monitored and recorded by a camera.

Apparatus. The SR system was set up in a room. The computer was located outside of the room, and the wiring was managed such that participants could not see any wires.

Stimuli. The room contained various visual objects, such as tables, chairs. None were closer than 1.5 m to the participants. Normal Question scene: During the recording session, an experimenter appeared from outside of the room. He moved to the front of the camera and asked questions to the camera as if talking to the participant then disappeared. After each question, the experimenter paused for a few seconds, which allowed time for a participant's response in the subsequent experience session. Fake Live scene: This scene was identical to the Normal Question scene except that the experimenter told participants that the current scene was live and asked participants whether they could tell that the given scene was live. Doppelgänger scene: During the recording session, the participants were invited into the room where the panoramic camera had already started recording. None of the participants paid attention to the camera because they were not informed that they would experience the movie taken by the camera in a later experience session (Fig. 2a). The experimenter and each participant had a 2~3 min conversation in front of the camera before the experimenter brought the participant out of the room. After the recording, experimenters removed the panoramic camera from the room and placed a height adjustable chair in the same location. Then the participant was invited into the room again for the experience session and sat in the chair.

Design and Procedure. Before setting up the experience module, the height of participants' eyes was adjusted to match the viewpoint of the live and recorded scenes. The procedure was also applied in Experiment II/III. First, the participants experienced a live scene, and later, the scene was manually switched to the Normal Question scene when they turned their heads (Fig. 2b). We examined whether substitution was successful with their ongoing reactions to the experimenter's questions and their reports after the whole experiment. The scene was then switched to the Doppelgänger scene, which lasted for approximately 2~3 minutes (Fig. 2c). After experiencing the Doppelgänger scene, ten of 21 participants experienced the additional Fake Live scene (Fig. 2d). The experimental session was finished when the scene was finally switched back to the live scene (Fig. 2e). The experimental module was removed after the participants had casual conversations with the experimenter.

Experiment II: Motion Parallax and the discrimination of scenes.
Participants. Ten adult volunteers (8 males and 2 females) served as paid participants (average age: 29.9 years). All had normal or corrected-to-normal vision. Six participants had previous experience of wearing a HMD.

Apparatus. The SR system was used. The headphones were disconnected because the scenes presented in the experiments were silent. The participants' head movements were monitored by the orientation sensor.

Stimuli. Live and recorded scenes contained the following content: a room with a white floor, black partitions (6.5 m from the participants) and a door. There was one red chair placed at three different locations in front of the partitions. The horizontal distance between the chair and each participant varied between 1.0 m, 2.5 m and 4.0 m. The recorded scenes were captured by the panoramic camera that had previously sat where the participant's head was located.

Design and Procedure. During the experiment, either a live or a recorded scene was pseudo-randomly selected and presented for 10 sec. Each scene was presented five times in one block. Thus, one block consisted of ten trials (example sequence: recorded, live, live, recorded, recorded, live, live, recorded, recorded, live) (Fig. 3c). During the inter-trial-interval (3 sec), a fixation target was presented at the HMD screen centre, and the participant was asked to focus on that target. In each trial, the participant reported whether a given scene was live or recorded by pressing a button on an input interface as soon as he/she became confident about the decision. The participant was instructed to pay attention to differences in the motion parallax associated with a red chair and its surroundings to discriminate the live and recorded scenes. One experimental session consisted of 6 blocks, as there were three different chair distances (1.0 m/2.5 m/4.0 m; see Fig. 3a) and two instructions regarding head movement (Fig. 3b). With the "Head Only" instruction, the participants were asked not to displace their body when they changed their head direction. With the "Head and Upper-body" instruction, the participants were asked to consciously displace their upper body (i.e., move their shoulders) and change their head orientation. The participants engaged in two experimental sessions following one training session. There was a resting period between blocks and sessions.

Experiment III: Head speed and detection of switching between scenes.
Participants. The same participants who participated in Experiment II were recruited for this study.

Apparatus. Identical to Experiment II.

Stimuli. Identical to Experiment II, except that the distance of the red chair was fixed at 2.5 m.

Design and Procedure. Participants were asked to turn their head intermittently at four speeds ("Motionless", "Slow", "Fast", and "Very Fast") without moving their torsos. The minimum target speed for each instruction was set at 0, 32, 64, or 96 deg/sec, respectively. Current head speed was measured using an orientation sensor and presented on the HMD display together with target speed. (Fig. 5a) When participants' head speed exceeded the minimum target speed, the switch between the live and recorded scenes was automatically executed, and the participants were asked to identify the switch as quickly as possible. The minimum duration of each scene presentation was randomly chosen from 5 to 15 sec so that the participant could not predict when the switch would occur (Fig. 5b). Within a chosen duration, the switch did not occur regardless of head speed until the duration limit had been met. After that time had passed, the switch was automatically executed when the head speed reached the threshold. There were 10 switches for each trial. One session consisted of four trials (one for each target speed). Three sessions were performed, and the speed conditions were randomised. If the participant pressed a button within 3 sec after the switch, it was categorised as a correct response.

Statistical Methods. For the response time in Experiment II and the correct detection rates in Experiment III, we applied repeated measure ANOVAs (analysis of variance). Jarque-Bera tests did not reject the hypotheses of normality of each data set (the smallest significance was 0.51 for the response time in Experiment II, and 0.13 for the correct detection rates in Experiment III, respectively). For the correct detection rates in Experiment II, we applied a Friedman's test, which is nonparametric, given that the rates were close to 100% and the requirements for parametric tests (i.e., normality and equality of variance) were not satisfied. Therefore, we only tested the null hypothesis that the correct rates were not modulated by the distance to the chair. This analysis satisfied our purpose because our concern here was the effect of the objects' distance on the correct detection rates.

1. Berrios, G. E. Disorientation states and psychiatry. *Compr Psychiatry* **23**, 479–491 (1982).
2. Hayashi, R., Watanabe, R., Mimura, M., Kato, M. & Katsumata, Y. A case of moyamoya disease presenting with geographic mislocation, person misidentification and fantastic confabulation. *No To Shinkei* **52**, 1091–1096 (2000).
3. Patterson, M. B. & Mack, J. L. Neuropsychological analysis of a case of reduplicative paramnesia. *J Clin Exp Neuropsychol* **7**, 111–121 (1985).
4. Pick, A. On reduplicative paramnesia. *Brain* **26**, 242–267 (1903).
5. Schnider, A. Spontaneous confabulation, reality monitoring, and the limbic system--a review. *Brain Res Brain Res Rev* **36**, 150–160 (2001).
6. Schnider, A. Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nat Rev Neurosci* **4**, 662–671 (2003).
7. Flavel, J. H. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist* **34**, 906–911 (1979).
8. Benjamin, A. S., Bjork, R. A. & Schwartz, B. L. The mismeasure of memory: when retrieval fluency is misleading as a metamnemonic index. *J Exp Psychol Gen* **127**, 55–68 (1998).
9. Koriat, A. in *Cambridge Handbook of Consciousness* (ed Zelazo, P. D. *et al*) 289–325 (2007).
10. Metcalfe, J. & Shimamura, A. *Metacognition: knowing about knowing* (MIT Press, 1994).
11. Nelson, T. O. *Metacognition: Core Readings* (Allyn and Bacon, 1992).
12. Nisbett, R. E. & Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. *Psychological Review* **84**, 231–259 (1977).
13. Hall, L., Johansson, P., Tarning, B., Sikstrom, S. & Deutgen, T. Magic at the marketplace: Choice blindness for the taste of jam and the smell of tea. *Cognition* **117**, 54–61 (2010).
14. Johansson, P., Hall, L., Sikstrom, S. & Olsson, A. Failure to detect mismatches between intention and outcome in a simple decision task. *Science* **310**, 116–119 (2005).
15. Martinez-Conde, S. & Macknik, S. L. Magic and the brain. *Sci Am* **299**, 72–79 (2008).
16. Johansson, P., Hall, L., Gulz, A., Haake, M. & Watanabe, K. Choice blindness and trust in the virtual world. *Technical report of IEICE: HIP* **107**, 83–86 (2007).
17. Bohil, C. J., Alicea, B. & Biocca, F. A. Virtual reality in neuroscience research and therapy. *Nat Rev Neurosci* **12**, 752–762 (2011).
18. Patterson, R., Winterbottom, M. D. & Pierce, B. J. Perceptual issues in the use of head-mounted visual displays. *Hum Factors* **48**, 555–573 (2006).
19. Beall, A. C., Loomis, J. M., Philbeck, J. W. & Fikes, T. G. Absolute Motion Parallax Weakly Determines Visual Scale In Real And Virtual Environments. *Proceedings of the Conference on Human Vision, Visual Processing, and Digital Display*, 288–297 (1995).
20. Adelstein, B. D., Li, L., Jerald, J. J. & Ellis, S. R. Suppressed Perception of Head-Referenced Image Motion during Head Movement. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **50**, 2678–2682 (2006).
21. Jay, C. & Hubbold, R. Amplifying Head Movements with Head-Mounted Displays. *Presence: Teleoperators and Virtual Environments* **12**, 268–276 (2003).
22. Jerald, J., Steinicke, F. & Whitton, M. Scene-Motion Thresholds Correlate with Angular Head Motions for Immersive Virtual Environments. *Proceedings of the 2009 Second International Conferences on Advances in Computer-Human Interactions*, 69–74 (2009).
23. Peck, T. C., Whitton, M. C. & Fuchs, H. Evaluation of Reorientation Techniques for Walking in Large Virtual Environments. *Virtual Reality Conference, 2008. VR '08. IEEE*, 121–127 (2008).

24. Coelho, D. C., Tichon, P. J., Hine, D. T., Wallis, D. G. & Riva, P. G. *Media Presence and Inner Presence: The Sense of Presence in Virtual Reality Technologies* (IOS Press, 2006).
25. Sanchez-Vives, M. V. & Slater, M. From presence to consciousness through virtual reality. *Nat Rev Neurosci* **6**, 332–339 (2005).
26. Slater, M., Lotto, B., Arnold, M. M. & Sanchez-Vives, M. V. How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de Psicología* **2**, 193–210 (2009).
27. Herrera, G., Jordan, R. & Vera, L. Agency and presence: a common dependence on subjectivity? *Presence: Tele-operators & Virtual Environments* **15**, 539–552 (2006).
28. Riva, G., Waterworth, J. A., Waterworth, E. L. & Mantovani, F. From intention to action: The role of presence. *New Ideas in Psychology* **29**, 24–37 (2011).
29. Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc Lond B Biol Sci* **364**, 3549–3557 (2009).
30. Costanza, E., Kunz, A. & Fjeld, M. Vol. 5440 *Lecture Notes in Computer Science* (eds Denis Lalanne & Jürg Kohlas) 47–68 (Springer Berlin / Heidelberg, 2009).
31. Herling, J. & Broll, W. Advanced Self-contained Object Removal for Realizing Real-time Diminished Reality in Unconstrained Environments. *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR' 2010)*, 207–212 (2010).
32. Ware, C., Arthur, K. & Booth, K. S. Fishtank Virtual Reality. *Proceedings of Conference on Human Factors in Computing Systems (INTERACT'93)*, 37–42 (1993).
33. Fiala, M. Pano-Presence for Teleoperation. *Proceedings of 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3798–3802 (2005).
34. Kato, N. *et al.* Omni-directional Stereoscopic Video Camera Platform for Visual Telexistence. *Transactions of the Virtual Reality Society of Japan* **13**, 353–362 (2008).
35. Corlett, P. R., Taylor, J. R., Wang, X. J., Fletcher, P. C. & Krystal, J. H. Toward a neurobiology of delusions. *Progress in neurobiology* **92**, 345–369 (2010).
36. Whitson, J. A. & Galinsky, A. D. Lacking control increases illusory pattern perception. *Science* **322**, 115–117 (2008).
37. Vuilleumier, P. Anosognosia: the neurology of beliefs and uncertainties. *Cortex* **40**, 9–17 (2004).
38. Coltheart, M., Langdon, R. & McKay, R. Schizophrenia and monothematic delusions. *Schizophr Bull* **33**, 642–647 (2007).
39. Vallar, G. & Ronchi, R. Somatoparaphrenia: a body delusion. A review of the neuropsychological literature. *Exp Brain Res* **192**, 533–551 (2009).
40. Emmelkamp, P. M., Bruynzeel, M., Drost, L. & van der Mast, C. A. Virtual reality treatment in acrophobia: a comparison with exposure in vivo. *Cyberpsychol Behav* **4**, 335–339 (2001).
41. Riva, G. Virtual reality in psychotherapy: review. *Cyberpsychol Behav* **8**, 220–240 (2005).
42. Rothbaum, B. O., Hodges, L. F., Ready, D., Graap, K. & Alarcon, R. D. Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *J Clin Psychiatry* **62**, 617–622 (2001).

Acknowledgements

Author contributions

KS and SW contributed equally to this work.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Suzuki, K., Wakisaka, S. & Fujii, N. Substitutional Reality System: A Novel Experimental Platform for Experiencing Alternative Reality. *Sci. Rep.* **2**, 459; DOI:10.1038/srep00459 (2012).